

SEQUENCE LISTING

5 <110> AGY Therapeutics
Melcher, Thorsten
10 Mueller, Sabine
Chin, Daniel

15 <120> USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZATION
OF BRAIN TUMORS

20 <130> 263/180 -- PEagleman -- AGY

25 <160> 35

30 <170> PatentIn version 3.1

35 <210> 1
<211> 3091
<212> DNA
<213> Homo sapiens

40 <220>
<221> gene

45 <222> (1)..(3091)
<223> PTP-zeta SM1 exon 9 variant

50 <220>
<221> VARSPLIC

55 <222> (1262)..(1262)

<223> Alternative splice site

5

<220>

<221> misc_feature

10

<222> (1273)..(3091)

<223> 3' Untranslated Region

15

<220>

<221> CDS

20

<222> (148)..(1272)

<223>

25

<400> 1

cacacatacg cacgcacgat ctcaacttoga tctatacact ggaggattaa aacaaacaaa 60

caaaaaaaaaac atttccttcg ctccccctcc ctctccactc tgagaagcag aggagccgca 120

cggcgagggg cgcgagaccg tctggaa atg cga atc cta aag cgt ttc ctc gct 174
Met Arg Ile Leu Lys Arg Phe Leu Ala
1 5

35 tgc att cag ctc ctc tgt gtt tgc cgc ctg gat tgg gct aat gga tac 222
Cys Ile Gln Leu Leu Cys Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr
10 15 20 25

40 tac aga caa cag aga aaa ctt gtt gaa gag att ggc tgg tcc tat aca 270
Tyr Arg Gln Gln Arg Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr
30 35 40

45 gga gca ctg aat caa aaa aat tgg gga aag aaa tat cca aca tgt aat 318
Gly Ala Leu Asn Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn
45 50 55

agc cca aaa caa tct cct atc aat att gat gaa gat ctt aca caa gta 366
Ser Pro Lys Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val
60 65 70

50 aat gtg aat ctt aag aaa ctt aaa ttt cag ggt tgg gat aaa aca tca 414
Asn Val Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser
75 80 85

55 ttg gaa aac aca ttc att cat aac act ggg aaa aca gtg gaa att aat 462
Leu Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn

5 90 95 100 105
 ctc act aat gac tac cgt gtc agc gga gga gtt tca gaa atg gtg ttt 510
 Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val Phe
 110 115 120
 10 aaa gca agc aag ata act ttt cac tgg gga aaa tgc aat atg tca tct 558
 Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met Ser Ser
 125 130 135
 gat gga tca gag cat agt tta gaa gga caa aaa ttt cca ctt gag atg 606
 Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro Leu Glu Met
 140 145 150
 15 caa atc tac tgc ttt gat gcg gac cga ttt tca agt ttt gag gaa gca 654
 Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser Phe Glu Glu Ala
 155 160 165
 20 gtc aaa gga aaa ggg aag tta aga gct tta tcc att ttg ttt gag gtt 702
 Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser Ile Leu Phe Glu Val
 170 175 180 185
 25 ggg aca gaa gaa aat ttg gat ttc aaa gcg att att gat gga gtc gaa 750
 Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala Ile Ile Asp Gly Val Glu
 190 195 200
 30 agt gtt agt cgt ttt ggg aag cag gct gct tta gat cca ttc ata ctg 798
 Ser Val Ser Arg Phe Gly Lys Gln Ala Ala Leu Asp Pro Phe Ile Leu
 205 210 215
 ttg aac ctt ctg cca aac tca act gac aag tat tac att tac aat ggc 846
 Leu Asn Leu Leu Pro Asn Ser Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly
 220 225 230
 35 tca ttg aca tct cct ccc tgc aca gac aca gtt gac tgg att gtt ttt 894
 Ser Leu Thr Ser Pro Pro Cys Thr Asp Thr Val Asp Trp Ile Val Phe
 235 240 245
 40 aaa gat aca gtt agc atc tct gaa agc cag ttg gct gtt ttt tgt gaa 942
 Lys Asp Thr Val Ser Ile Ser Glu Ser Gln Leu Ala Val Phe Cys Glu
 250 255 260 265
 45 gtt ctt aca atg caa caa tct ggt tat gtc atg ctg atg gac tac tta 990
 Val Leu Thr Met Gln Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu
 270 275 280
 caa aac aat ttt cga gag caa cag tac aag ttc tct aga cag gtg ttt 1038
 Gln Asn Asn Phe Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe
 285 290 295
 50 tcc tca tac act gga aag gaa gag att cat gaa gca gtt tgt agt tca 1086
 Ser Ser Tyr Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser
 300 305 310
 55 gaa cca gaa aat gtt cag gct gac cca gag aat tat acc agc ctt ctt 1134
 Glu Pro Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu

	315	320	325	
5	gtt aca tgg gaa aga cct cga gtc gtt tat gat acc atg att gag aag Val Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys 330 335 340 345			1182
10	ttt gca gtt ttg tac cag cag ttg gat gga gag gac caa acc aag cat Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys His 350 355 360			1230
	gaa ttt ttg aca gat ggc tat caa gac ttg gta act ata tga Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Val Thr Ile 365 370			1272
15	tcagttgttt tacatagggg aacattataa tttaatttcc aaggttaagaa cttacaaatg			1332
20	gttgatatatt attttctctc attactttta gactttatgt gaaggtgggg taggctgagt			1392
	attttttaaat ttaaaaaaaaa atttttaatt agaagctata ctaaattatg tttaaagtta			1452
25	catttaatta aatggatata ataactttgc caacaataac actatagagt agatacatat			1512
	gacttatgaa ctggagatca tttagtgtgg cttttcttaa gatttcagtt gtagaatagt			1572
30	gccagaatct cagtgccttg atacatttta tattgtgtct tccattacgc tatatcagca			1632
	caggaaaagt agagtagggg acatacaagt cctcttttgtt gcaccaaaaa attttcagat			1692
	aacagctggg aagtcatgat tgggtcagaa ctttggggat gtaagaaaac atttcttaca			1752
	aaaagatcca cccctgcctc cctccaccag cgcatgcaa taaagtacag attccctttg			1812
	tggcctgagc atgtcagtat taaactttgc tctggtaggg aagtgttggc catagattag			1872
35	gggtgtagttg acaaaccttc atctggatgt aggtccagaa agtccccact gcagggttaa			1932
	ggacactgga ctctgcactc aggcacctag agtcctgcaa gtcctgggaa cctgcattta			1992
40	aataaaaatg cactattaat tatgtttcat atcatgtgga caaaatggat aaaattttag			2052
	taacctttta attcagttgc ctggaatatg gagacacaat gacctgggaa aatcgtgaaa			2112
	taaatagtaa taaaaatggt tatttcataa ttacgtgaag aagataattc tattactggt			2172
45	cttgcatata tattgtcaag aaaaagagat aacttagttg ttactttttt cacattgctc			2232
	cttgtttgca aatgcccccc atttatttgt ctaaaatatt aatttttagt ttgtagtact			2292
50	aatttatgaa tttgatgagt tctggctaaa aatgaaactt cctgaaacta aatctgattt			2352
	ttaaaaagca aaaaaaaaaa aaaagcctag ctttccagtt cttcataatt cacaaatacc			2412
	acaagtttaa ctaageaaca ttgcataaac ttttccttag gttaataaaa tagaagtatt			2472
55	ttccacggac cagggagaaa aagttttcta ggaaagatac ctagtgtgtt ggtagtccta			2532

5

25

35

40

45

50

55

OC-94692.1

50 55 60

5 Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu
65 70 75 80

10 Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His
85 90 95

15 Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val
100 105 110

20 Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe
115 120 125

25 His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu
130 135 140

30 Glu Gly Gln Lys Phe Pro Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala
145 150 155 160

35 Asp Arg Phe Ser Ser Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu
165 170 175

40 Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp
180 185 190

45 Phe Lys Ala Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys
195 200 205

50 Gln Ala Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser
210 215 220

55 Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
225 230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser
245 250 255

Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser
260 265 270

Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln

275

280

285

5

Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu
290 295 300

10

Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala
305 310 315 320

15

Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg
325 330 335

20

Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln
340 345 350

25

Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr
355 360 365

Gln Asp Leu Val Thr Ile
370

30

<210> 3

<211> 8058

<212> DNA

<213> Homo sapiens

35

<220>

40

<221> gene

<222> (1)..(8058)

45

<223> PTP-zeta SM2 exon 23a variant

<220>

50

<221> VARSPLIC

<222> (6229)..(6345)

55

<223> Exon 23a

<220>

<221> misc_feature

<222> (7210)..(8058)

<223> 3' Untranslated Region

<400> 3

cacacatacg cagcgacgat ctccacttga tctatacact ggaggattaa aacaaacaaa 60

caaaaaaaaaac atttccttcg ctccccctcc ctctccactc tgagaagcag aggagccgca 120

cggcgagggg cgcgagaccg tctggaaatg cgaatcctaa agcgtttcct cgcttgcaatt 180

cagctcctct gtgtttgccg cctggattgg gctaattgat actacagaca acagagaaaa 240

cttggtgaag agattggctg gtcctataca ggagcactga atcaaaaaaaaa ttggggaaaag 300

aaatatccaa catgtaatag cccaaaacaa tctcctatca atattgatga agatcttaca 360

caagtaaatg tgaatcttaa gaaacttaaa ttccagggtt gggataaaac atcattggaa 420

aacacattca ttcataacac tgggaaaaca gtggaaatta atctcactaa tgactaccgt 480

gtcagcggag gagtttcaga aatgggtgtt aaagcaagca agataacttt tccactgggga 540

aatgcaata tgtcatctga tggatcagag catagttagt aaggacaaaa atttcactt 600

gagatgcaaa tctactgctt tgatgcggac cgattttcaa gttttgagga agcagtcaaa 660

ggaaaaggga agttaagagc tttatccatt ttgtttgagg ttgggacaga agaaaatttg 720

gatttcaaag cgattattga tggagtcgaa agtggttagtc gttttgggaa gcaggctgct 780

ttagatccat tcatactgtt gaaccttctg ccaaactcaa ctgacaagta ttacatttac 840

aatggctcat tgacatctcc tccctgcaca gacacagttg actggattgt ttttaaagat 900

acagttagca tctctgaaag ccagttggct gttttttgtg aagttcttac aatgcaacaa 960

tctggttatg tcatgctgat ggactactta caaaacaatt ttcgagagca acagtacaag 1020

ttctctagac aggtgttttc ctccatacact ggaaaggaag agattcatga agcagtttgt 1080

agttcagaac cagaaaatgt tcaggctgac ccagagaatt ataccagcct tcttggtaca 1140

tgggaaagac ctcgagtcgt ttatgatacc atgattgaga agtttgcagt tttgtaccag 1200

cagttggatg gagaggacca aaccaagcat gaatttttga cagatggcta tcaagacttg 1260

ggtgctattc tcaataatth gctacccaat atgagttatg ttcttcagat agtagccata 1320

50

5 10 15 20 25 30 35 40 45 50 55

gattctgtgg gtgtaactta tcagggttcc ttatttagcg gccctagcca tataccaata 3060

cctaagtctt cgtaataaac cccaactgca tcattactgc agcctactca tgccctctct 3120

ggtgatgggg aatggtctgg agcctcttct gatagtgaat ttcttttacc tgacacagat 3180

gggctgacag cccttaacat ttcttcacct gtttctgtag ctgaatttac atatacaaca 3240

tctgtgtttg gtgatgataa taaggcgctt tctaaaagtg aaataatata tggaaatgag 3300

actgaactgc aaattccttc tttcaatgag atgggttacc cttctgaaag cacagtcagt 3360

cccaacatgt atgataatgt aaataagttg aatgcgtctt tacaagaaac ctctgtttcc 3420

atttctagca ccaagggcat gtttccaggg tcccttgctc ataccaccac taagggtttt 3480

gatcatgaga ttagtcaagt tccagaaaat aacttttcag ttcaacctac acatactgtc 3540

tctcaagcat ctggtgacac ttctgtttaa cctgtgctta gtgcaaactc agagccagca 3600

tcctctgacc ctgcttctag tgaaatgtta tctccttcaa ctcagctctt attttatgag 3660

acctcagctt ctttttagtac tgaagtattg ctacaacctt cttttcaggc ttctgatggt 3720

gacaccttgc ttaaaactgt tcttccagct gtgccagtg atccaatatt ggttgaaacc 3780

cccaaagtgt ataaaattag ttctacaatg ttgcatctca ttgtatcaaa ttctgcttca 3840

agtgaaaaca tgctgcactc tacatctgta ccagtttttg atgtgtcgcc tacttctcat 3900

atgcactctg cttcacttca aggtttgacc atttcctatg caagtgagaa atatgaacca 3960

gttttgtaa aaagtgaag ttcccaccaa gtggtacctt ctttgtagag taatgatgag 4020

ttgttccaaa cggccaattt ggagattaac caggcccatc ccccaaaagg aaggcatgta 4080

tttgctacac ctgtttttatc aattgatgaa ccattaaata cactaataaa taagcttata 4140

cattccgatg aaattttaac ctccaccaa agttctgtta ctggttaagg atttgctggt 4200

attccaacag ttgcttctga tacatttgta tctactgatc attctgttcc tataggaaat 4260

gggcatgttg ccattacagc tgtttctccc cacagagatg gttctgtaac ctcaacaaag 4320

ttgctgtttc cttctaaggc aacttctgag ctgagtcata gtgccaaatc tgatgccggt 4380

ttagtgggtg gtggtgaaga tggtagact gatgatgatg gtgatgatga tgatgacaga 4440

gatagtgatg gcttatccat tcataagtgt atgtcatgct catcctatag agaatcacag 4500

gaaaaggtaa tgaatgattc agacaccac gaaaacagtc ttatggatca gaataatcca 4560

atctcact cactatctga gaattctgaa gaagataata gagtcacaag tgtatcctca 4620

gacagtcaaa ctggtatgga cagaagtcct ggtaaatac catcagcaaa tgggctatcc 4680

5
10
15
20
25
30
35
40
45
50
55

caaaagcaca atgatggaaa agaggaaaat gacattcaga ctggtagtgc tctgcttcct 4740
ctcagccctg aatctaaagc atgggcagtt ctgacaagtg atgaagaaaag tggatcaggg 4800
caaggtacct cagatagcct taatgagaat gagacttcca cagatttcag ttttgcagac 4860
actaatgaaa aagatgctga tgggatcctg gcagcagggtg actcagaaaat aactcctgga 4920
ttcccacagt ccccaacatc atctgttact agcgagaact cagaagtgtt ccacgtttca 4980
gaggcagagg ccagtaatat tagccatgag tctcgtattg gtctagctga ggggttgga 5040
tccgagaaga aggcagttat accccttggtg atcgtgtcag ccctgacttt tatctgtcta 5100
gtggttcttg tgggtattct catctactgg aggaaatgct tccagactgc acacttttac 5160
ttagaggaca gtacatcccc tagagttata tccacacctc caacacctat ctttccaatt 5220
tcagatgatg tcggagcaat tccaataaag cactttccaa agcatgttg agatttacat 5280
gcaagtagtg ggtttactga agaatttgag aactgaaag agttttacca ggaagtgcag 5340
agctgtactg ttgacttagg tattacagca gacagctcca accaccaga caacaagcac 5400
aagaatcgat acataaatat cgttgccctat gatcatagca gggtaagct agcacagctt 5460
gctgaaaagg atggcaaact gactgattat atcaatgcc aattatgttg tggctacaac 5520
agaccaaag cttatattgc tgcccaaggc cactgaaat ccacagctga agatttctgg 5580
agaatgatat gggaacataa tgtggaagtt attgtcatga taacaaacct cgtggagaaa 5640
ggaaggagaa aatgtgatca gtactggcct gccgatggga gtgaggagta cggaacttt 5700
ctggtcactc agaagagtgt gcaagtgctt gcctattata ctgtgaggaa ttttactcta 5760
agaaacacaa aaataaaaaa gggctcccag aaaggaagac ccagtggacg tgtggtcaca 5820
cagtatcact acacgcagtg gcctgacatg ggagtaccag agtactccct gccagtgtg 5880
acctttgtga gaaaggcagc ctatgccaaag cgccatgcag tggggcctgt tgtcgtccac 5940
tgcaagtgtg gagttggaag aacaggcaca tatattgtgc tagacagtat gttgcagcag 6000
attcaacacg aaggaactgt caacatattt ggcttcttaa aacacatccg ttcacaaaga 6060
aattatttgg taaaaactga ggagcaatat gtcttcattc atgatacact ggttgaggcc 6120
atacttagta aagaaactga ggtgctggac agtcatattc atgcctatgt taatgcactc 6180
ctcattcctg gaccagcagg caaaacaaag ctagagaaac aattccaggg tctcactctg 6240
tcaccaggc tggagtgcag aggcacaatc tcggctcact gcaaccttcc tctccctggc 6300
ttaactgatc ctctacctc agcctcccga gtggctggga ctatactcct gagccagtca 6360

5
10
15
20
25
30
35
40
45
50
55

aatatacagc agagtgacta ttctgcagcc ctaaagcaat gcaacagggg aaagaatcga 6420
acttcttcta tcatccctgt ggaaagatca aggggttgga tttcatccct gagtggagaa 6480
ggcacagact acatcaatgc ctctatatatc atgggctatt accagagcaa tgaattcatc 6540
attaccagc accctctcct tcataccatc aaggatttct ggaggatgat atgggacat 6600
aatgccccaa tgggtggttat gattcctgat ggccaaaaca tggcagaaga tgaatttggt 6660
tactggccaa ataaagatga gcctataaat tgtgagagct ttaaggtcac tcttatggct 6720
gaagaacaca aatgtctatc taatgaggaa aaacttataa ttcaggactt tatcttagaa 6780
gctacacagg atgattatgt acttgaagtg aggcactttc agtgtcctaa atggccaaat 6840
ccagatagcc ccattagtaa aacttttgaa cttataagtg ttataaaaga agaagctgcc 6900
aatagggatg ggcctatgat tgttcatgat gagcatggag gagtgcggc aggaactttc 6960
tgtgctctga caacccttat gcaccaacta gaaaaagaaa attccgtgga tgtttaccag 7020
gtagccaaga tgatcaatct gatgaggcca ggagtcttg ctgacattga gcagtatcag 7080
tttctctaca aagtgatcct cagccttggt agcacaaggc aggaagagaa tccatccacc 7140
tctctggaca gtaatggtgc agcattgcct gatggaaata tagctgagag cttagagtct 7200
ttagtttaac acagaaaggg gtgggggggac tcacatctga gcattgtttt cctcttccta 7260
aaattaggca ggaaaatcag tctagtcttg ttatctgttg atttcccatc acctgacagt 7320
aactttcatg acataggatt ctgccgcaa atttatatca ttaacaatgt gtgccttttt 7380
gcaagacttg taatttactt attatgtttg aactaaaatg attgaatttt acagtatttc 7440
taagaatgga attgtggtat ttttttctgt attgatttta acagaaaatt tcaatttata 7500
gaggtttaga attccaaact acagaaaatg tttgttttta gtgtcaaatt tttagctgta 7560
tttgtagcaa ttatcagggt tgctagaaat ataactttta atacagtagc ctgtaaataa 7620
aacactcttc catatgatat tcaacatttt acaactgcag tattcaccta aagtagaaat 7680
aatctgttac ttattgtaaa tactgcccta gtgtctccat ggaccaaatt tatatttata 7740
attgtagatt tttatatttt actactgagt caagttttct agttctgtgt aattgtttag 7800
tttaatgacg tagttcatta gctggtctta ctctaccagt tttctgacat tgtattgtgt 7860
tacctaagtc attaactttg tttcagcatg taattttaac ttttgtggaa aatagaaata 7920
ccttcatttt gaaagaagtt tttatgagaa taacacctta ccaaacattg ttcaaattgg 7980
ttttatccaa ggaattgcaa aaataaatat aaatattgcc attaaaaaaa aaaaaaaaaa 8040

aaaaaaaaaa aaaaaaaaa

8058

5 <210> 4

<211> 2353

<212> PRT

10 <213> Homo sapiens

15 <220>

<221> gene

<222> (1) .. (2353)

<223> PTP-zeta SM2 23a exon variant

20 <400> 4

Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys Val
1 5 10 15

Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg Lys Leu
20 25 30

35 Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn Gln Lys Asn
35 40 45

40 Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys Gln Ser Pro Ile
50 55 60

Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu
65 70 75 80

45 Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His
85 90 95

50 Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val
100 105 110

55 Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe
115 120 125

5 His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu
130 135 140

10 Glu Gly Gln Lys Phe Pro Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala
145 150 155 160

15 Asp Arg Phe Ser Ser Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu
165 170 175

20 Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp
180 185 190

25 Phe Lys Ala Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys
195 200 205

30 Gln Ala Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser
210 215 220

35 Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
225 230 235 240

40 Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser
245 250 255

45 Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser
260 265 270

50 Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln
275 280 285

55 Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu
290 295 300

60 Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala
305 310 315 320

65 Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg
325 330 335

70 Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln
340 345 350

5 Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr
355 360 365

10 Gln Asp Leu Gly Ala Ile Leu Asn Asn Leu Leu Pro Asn Met Ser Tyr
370 375 380

15 Val Leu Gln Ile Val Ala Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr
385 390 395 400

20 Ser Asp Gln Leu Ile Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp
405 410 415

25 Leu Phe Pro Glu Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu
420 425 430

30 Glu Gly Lys Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp
435 440 445

35 Ser Ala Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr
450 455 460

40 Thr His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn
465 470 475 480

45 Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val Pro
485 490 495

50 Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu Ala Thr
500 505 510

55 Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu Leu Pro Pro
515 520 525

60 His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp Gly Ser Lys Thr
530 535 540

65 Val Leu Arg Ser Pro His Met Asn Leu Ser Gly Thr Ala Glu Ser Leu
545 550 555 560

70 Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu Glu Ser Leu Leu Thr Ser
565 570 575

5

10

15

20

25

30

35

40

45

50

55

5 Pro Ser Val Asp Val Ser Phe Glu Ser Ile Leu Ser Ser Tyr Asp Gly
805 810 815

10 Ala Pro Leu Leu Pro Phe Ser Ser Ala Ser Phe Ser Ser Glu Leu Phe
820 825 830

15 Arg His Leu His Thr Val Ser Gln Ile Leu Pro Gln Val Thr Ser Ala
835 840 845

20 Thr Glu Ser Asp Lys Val Pro Leu His Ala Ser Leu Pro Val Ala Gly
850 855 860

25 Gly Asp Leu Leu Leu Glu Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu
865 870 875 880

30 Ser Thr Thr His Ala Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser
885 890 895

35 Gly Val Leu Tyr Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser
900 905 910

40 Ser Asp Ala Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr
915 920 925

45 Ala Leu Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr
930 935 940

50 Ser Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly
945 950 955 960

55 Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser Leu
965 970 975

60 Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu Ser Gly
980 985 990

65 Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe Leu Leu Pro
995 1000 1005

70 Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser Pro Val Ser
1010 1015 1020

5 Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly Asp Asp Asn
1025 1030 1035

10 Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn Glu Thr Glu
1040 1045 1050

15 Leu Gln Ile Pro Ser Phe Asn Glu Met Val Tyr Pro Ser Glu Ser
1055 1060 1065

20 Thr Val Met Pro Asn Met Tyr Asp Asn Val Asn Lys Leu Asn Ala
1070 1075 1080

25 Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr Lys Gly Met
1085 1090 1095

30 Phe Pro Gly Ser Leu Ala His Thr Thr Thr Lys Val Phe Asp His
1100 1105 1110

35 Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val Gln Pro Thr
1115 1120 1125

40 His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu Lys Pro Val
1130 1135 1140

45 Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro Ala Ser Ser
1145 1150 1155

50 Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr Glu Thr Ser
1160 1165 1170

55 Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser Phe Gln Ala
1175 1180 1185

Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro Ala Val Pro
1190 1195 1200

Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp Lys Ile Ser
1205 1210 1215

Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala Ser Ser Glu
1220 1225 1230

5
Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp Val Ser Pro
1235 1240 1245

10
Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu Thr Ile Ser
1250 1255 1260

15
Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys Ser Glu Ser
1265 1270 1275

20
Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp Glu Leu Phe
1280 1285 1290

25
Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro Pro Lys Gly
1295 1300 1305

30
Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp Glu Pro Leu
1310 1315 1320

35
Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu Ile Leu Thr
1325 1330 1335

40
Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala Gly Ile Pro
1340 1345 1350

45
Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His Ser Val Pro
1355 1360 1365

50
Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser Pro His Arg
1370 1375 1380

55
Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro Ser Lys Ala
1385 1390 1395

60
Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala Gly Leu Val
1400 1405 1410

65
Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly Asp Asp Asp
1415 1420 1425

70
Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys Cys Met Ser
1430 1435 1440

5 Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met Asn Asp Ser
1445 1450 1455

10 Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn Pro Ile Ser
1460 1465 1470

15 Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg Val Thr Ser
1475 1480 1485

20 Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser Pro Gly Lys
1490 1495 1500

25 Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn Asp Gly Lys
1505 1510 1515

30 Glu Glu Asn Asp Ile Gln Thr Gly Ser Ala Leu Leu Pro Leu Ser
1520 1525 1530

35 Pro Glu Ser Lys Ala Trp Ala Val Leu Thr Ser Asp Glu Glu Ser
1535 1540 1545

40 Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu Asn Glu Thr
1550 1555 1560

45 Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys Asp Ala Asp
1565 1570 1575

50 Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro Gly Phe Pro
1580 1585 1590

55 Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser Glu Val Phe
1595 1600 1605

60 His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His Glu Ser Arg
1610 1615 1620

65 Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys Ala Val Ile
1625 1630 1635

70 Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys Leu Val Val
1640 1645 1650

5 Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe Gln Thr Ala
1655 1660 1665

10 His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val Ile Ser Thr
1670 1675 1680

15 Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val Gly Ala Ile
1685 1690 1695

20 Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu His Ala Ser
1700 1705 1710

25 Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu Phe Tyr Gln
1715 1720 1725

30 Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr Ala Asp Ser
1730 1735 1740

35 Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr Ile Asn Ile
1745 1750 1755

40 Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln Leu Ala Glu
1760 1765 1770

45 Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn Tyr Val Asp
1775 1780 1785

50 Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln Gly Pro Leu
1790 1795 1800

55 Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp Glu His Asn
1805 1810 1815

60 Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu Lys Gly Arg
1820 1825 1830

65 Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser Glu Glu Tyr
1835 1840 1845

70 Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val Leu Ala Tyr
1850 1855 1860

5 Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys Ile Lys Lys
1865 1870 1875

10 Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val Thr Gln Tyr
1880 1885 1890

15 His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu Tyr Ser Leu
1895 1900 1905

20 Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala Lys Arg His
1910 1915 1920

25 Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly Val Gly Arg
1925 1930 1935

30 Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln Gln Ile Gln
1940 1945 1950

35 His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys His Ile Arg
1955 1960 1965

40 Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln Tyr Val Phe
1970 1975 1980

45 Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys Glu Thr Glu
1985 1990 1995

50 Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala Leu Leu Ile
2000 2005 2010

55 Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln Phe Gln Gly
2015 2020 2025

60 Leu Thr Leu Ser Pro Arg Leu Glu Cys Arg Gly Thr Ile Ser Ala
2030 2035 2040

65 His Cys Asn Leu Pro Leu Pro Gly Leu Thr Asp Pro Pro Thr Ser
2045 2050 2055

70 Ala Ser Arg Val Ala Gly Thr Ile Leu Leu Ser Gln Ser Asn Ile
2060 2065 2070

5 Gln Gln Ser Asp Tyr Ser Ala Ala Leu Lys Gln Cys Asn Arg Glu
2075 2080 2085

10 Lys Asn Arg Thr Ser Ser Ile Ile Pro Val Glu Arg Ser Arg Val
2090 2095 2100

15 Gly Ile Ser Ser Leu Ser Gly Glu Gly Thr Asp Tyr Ile Asn Ala
2105 2110 2115

20 Ser Tyr Ile Met Gly Tyr Tyr Gln Ser Asn Glu Phe Ile Ile Thr
2120 2125 2130

25 Gln His Pro Leu Leu His Thr Ile Lys Asp Phe Trp Arg Met Ile
2135 2140 2145

30 Trp Asp His Asn Ala Gln Leu Val Val Met Ile Pro Asp Gly Gln
2150 2155 2160

35 Asn Met Ala Glu Asp Glu Phe Val Tyr Trp Pro Asn Lys Asp Glu
2165 2170 2175

40 Pro Ile Asn Cys Glu Ser Phe Lys Val Thr Leu Met Ala Glu Glu
2180 2185 2190

45 His Lys Cys Leu Ser Asn Glu Glu Lys Leu Ile Ile Gln Asp Phe
2195 2200 2205

50 Ile Leu Glu Ala Thr Gln Asp Asp Tyr Val Leu Glu Val Arg His
2210 2215 2220

55 Phe Gln Cys Pro Lys Trp Pro Asn Pro Asp Ser Pro Ile Ser Lys
2225 2230 2235

Thr Phe Glu Leu Ile Ser Val Ile Lys Glu Glu Ala Ala Asn Arg
2240 2245 2250

60 Asp Gly Pro Met Ile Val His Asp Glu His Gly Gly Val Thr Ala
2255 2260 2265

65 Gly Thr Phe Cys Ala Leu Thr Thr Leu Met His Gln Leu Glu Lys
2270 2275 2280

Glu Asn Ser Val Asp Val Tyr Gln Val Ala Lys Met Ile Asn Leu
2285 2290 2295

5

Met Arg Pro Gly Val Phe Ala Asp Ile Glu Gln Tyr Gln Phe Leu
2300 2305 2310

10

Tyr Lys Val Ile Leu Ser Leu Val Ser Thr Arg Gln Glu Glu Asn
2315 2320 2325

15

Pro Ser Thr Ser Leu Asp Ser Asn Gly Ala Ala Leu Pro Asp Gly
2330 2335 2340

Asn Ile Ala Glu Ser Leu Glu Ser Leu Val
2345 2350

20

<210> 5

<211> 7941

<212> DNA

<213> Homo sapiens

30

<220>

35

<221> CDS

<222> (148)..(7092)

<223>

40

<220>

45

<221> gene

<222> (1)..(7941)

<223> PTP-zeta

50

<400> 5

cacacatacg cacgcacgat ctcacttcga tctatacact ggaggattaa aacaaacaaa 60

55

caaaaaaaaaac atttccttcg ctccccctcc ctctccactc tgagaagcag aggagccgca 120

5
10
15
20
25
30
35
40
45
50
55

cggcgagggg ccgcagaccg tctggaa atg cga atc cta aag cgt ttc ctc gct 174
Met Arg Ile Leu Lys Arg Phe Leu Ala
1 5

tgc att cag ctc ctc tgt gtt tgc cgc ctg gat tgg gct aat gga tac 222
Cys Ile Gln Leu Leu Cys Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr
10 15 20 25

tac aga caa cag aga aaa ctt gtt gaa gag att ggc tgg tcc tat aca 270
Tyr Arg Gln Gln Arg Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr
30 35 40

gga gca ctg aat caa aaa aat tgg gga aag aaa tat cca aca tgt aat 318
Gly Ala Leu Asn Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn
45 50 55

agc cca aaa caa tct cct atc aat att gat gaa gat ctt aca caa gta 366
Ser Pro Lys Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val
60 65 70

aat gtg aat ctt aag aaa ctt aaa ttt cag ggt tgg gat aaa aca tca 414
Asn Val Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser
75 80 85

ttg gaa aac aca ttc att cat aac act ggg aaa aca gtg gaa att aat 462
Leu Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn
90 95 100 105

ctc act aat gac tac cgt gtc agc gga gga gtt tca gaa atg gtg ttt 510
Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val Phe
110 115 120

aaa gca agc aag ata act ttt cac tgg gga aaa tgc aat atg tca tct 558
Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met Ser Ser
125 130 135

gat gga tca gag cat agt tta gaa gga caa aaa ttt cca ctt gag atg 606
Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro Leu Glu Met
140 145 150

caa atc tac tgc ttt gat gcg gac cga ttt tca agt ttt gag gaa gca 654
Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser Phe Glu Glu Ala
155 160 165

gtc aaa gga aaa ggg aag tta aga gct tta tcc att ttg ttt gag gtt 702
Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser Ile Leu Phe Glu Val
170 175 180 185

ggg aca gaa gaa aat ttg gat ttc aaa gcg att att gat gga gtc gaa 750
Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala Ile Ile Asp Gly Val Glu
190 195 200

agt gtt agt cgt ttt ggg aag cag gct gct tta gat cca ttc ata ctg 798
Ser Val Ser Arg Phe Gly Lys Gln Ala Ala Leu Asp Pro Phe Ile Leu
205 210 215

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ttg aac ctt ctg cca aac tca act gac aag tat tac att tac aat ggc 846
 Leu Asn Leu Leu Pro Asn Ser Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly
 220 225 230
 tca ttg aca tct cct ccc tgc aca gac aca gtt gac tgg att gtt ttt 894
 Ser Leu Thr Ser Pro Pro Cys Thr Asp Thr Val Asp Trp Ile Val Phe
 235 240 245
 aaa gat aca gtt agc atc tct gaa agc cag ttg gct gtt ttt tgt gaa 942
 Lys Asp Thr Val Ser Ile Ser Glu Ser Gln Leu Ala Val Phe Cys Glu
 250 255 260 265
 gtt ctt aca atg caa caa tct ggt tat gtc atg ctg atg gac tac tta 990
 Val Leu Thr Met Gln Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu
 270 275 280
 caa aac aat ttt cga gag caa cag tac aag ttc tct aga cag gtg ttt 1038
 Gln Asn Asn Phe Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe
 285 290 295
 tcc tca tac act gga aag gaa gag att cat gaa gca gtt tgt agt tca 1086
 Ser Ser Tyr Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser
 300 305 310
 gaa cca gaa aat gtt cag gct gac cca gag aat tat acc agc ctt ctt 1134
 Glu Pro Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu
 315 320 325
 gtt aca tgg gaa aga cct cga gtc gtt tat gat acc atg att gag aag 1182
 Val Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys
 330 335 340 345
 ttt gca gtt ttg tac cag cag ttg gat gga gag gac caa acc aag cat 1230
 Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys His
 350 355 360
 gaa ttt ttg aca gat ggc tat caa gac ttg ggt gct att ctc aat aat 1278
 Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Gly Ala Ile Leu Asn Asn
 365 370 375
 ttg cta ccc aat atg agt tat gtt ctt cag ata gta gcc ata tgc act 1326
 Leu Leu Pro Asn Met Ser Tyr Val Leu Gln Ile Val Ala Ile Cys Thr
 380 385 390
 aat ggc tta tat gga aaa tac agc gac caa ctg att gtc gac atg cct 1374
 Asn Gly Leu Tyr Gly Lys Tyr Ser Asp Gln Leu Ile Val Asp Met Pro
 395 400 405
 act gat aat cct gaa ctt gat ctt ttc cct gaa tta att gga act gaa 1422
 Thr Asp Asn Pro Glu Leu Asp Leu Phe Pro Glu Leu Ile Gly Thr Glu
 410 415 420 425
 gaa ata atc aag gag gag gaa gag gga aaa gac att gaa gaa ggc gct 1470
 Glu Ile Ile Lys Glu Glu Glu Glu Gly Lys Asp Ile Glu Glu Gly Ala
 430 435 440

att gtg aat cct ggt aga gac agt gct aca aac caa atc agg aaa aag 1518
Ile Val Asn Pro Gly Arg Asp Ser Ala Thr Asn Gln Ile Arg Lys Lys
445 450 455

5 gaa ccc cag att tct acc aca aca cac tac aat cgc ata ggg acg aaa 1566
Glu Pro Gln Ile Ser Thr Thr Thr His Tyr Asn Arg Ile Gly Thr Lys
460 465 470

10 tac aat gaa gcc aag act aac cga tcc cca aca aga gga agt gaa ttc 1614
Tyr Asn Glu Ala Lys Thr Asn Arg Ser Pro Thr Arg Gly Ser Glu Phe
475 480 485

15 tct gga aag ggt gat gtt ccc aat aca tct tta aat tcc act tcc caa 1662
Ser Gly Lys Gly Asp Val Pro Asn Thr Ser Leu Asn Ser Thr Ser Gln
490 495 500 505

20 cca gtc act aaa tta gcc aca gaa aaa gat att tcc ttg act tct cag 1710
Pro Val Thr Lys Leu Ala Thr Glu Lys Asp Ile Ser Leu Thr Ser Gln
510 515 520

25 act gtg act gaa ctg cca cct cac act gtg gaa ggt act tca gcc tct 1758
Thr Val Thr Glu Leu Pro Pro His Thr Val Glu Gly Thr Ser Ala Ser
525 530 535

30 tta aat gat ggc tct aaa act gtt ctt aga tct cca cat atg aac ttg 1806
Leu Asn Asp Gly Ser Lys Thr Val Leu Arg Ser Pro His Met Asn Leu
540 545 550

35 tcg ggg act gca gaa tcc tta aat aca gtt tct ata aca gaa tat gag 1854
Ser Gly Thr Ala Glu Ser Leu Asn Thr Val Ser Ile Thr Glu Tyr Glu
555 560 565

40 gag gag agt tta ttg acc agt ttc aag ctt gat act gga gct gaa gat 1902
Glu Glu Ser Leu Leu Thr Ser Phe Lys Leu Asp Thr Gly Ala Glu Asp
570 575 580 585

45 tct tca ggc tcc agt ccc gca act tct gct atc cca ttc atc tct gag 1950
Ser Ser Gly Ser Ser Pro Ala Thr Ser Ala Ile Pro Phe Ile Ser Glu
590 595 600

50 aac ata tcc caa ggg tat ata ttt tcc tcc gaa aac cca gag aca ata 1998
Asn Ile Ser Gln Gly Tyr Ile Phe Ser Ser Glu Asn Pro Glu Thr Ile
605 610 615

55 aca tat gat gtc ctt ata cca gaa tct gct aga aat gct tcc gaa gat 2046
Thr Tyr Asp Val Leu Ile Pro Glu Ser Ala Arg Asn Ala Ser Glu Asp
620 625 630

50 tca act tca tca ggt tca gaa gaa tca cta aag gat cct tct atg gag 2094
Ser Thr Ser Ser Gly Ser Glu Glu Ser Leu Lys Asp Pro Ser Met Glu
635 640 645

55 gga aat gtg tgg ttt cct agc tct aca gac ata aca gca cag ccc gat 2142
Gly Asn Val Trp Phe Pro Ser Ser Thr Asp Ile Thr Ala Gln Pro Asp
650 655 660 665

5
gtt gga tca ggc aga gag agc ttt ctc cag act aat tac act gag ata 2190
Val Gly Ser Gly Arg Glu Ser Phe Leu Gln Thr Asn Tyr Thr Glu Ile
670 675 680

10
cgt gtt gat gaa tct gag aag aca acc aag tcc ttt tct gca ggc cca 2238
Arg Val Asp Glu Ser Glu Lys Thr Thr Lys Ser Phe Ser Ala Gly Pro
685 690 695

15
gtg atg tca cag ggt ccc tca gtt aca gat ctg gaa atg cca cat tat 2286
Val Met Ser Gln Gly Pro Ser Val Thr Asp Leu Glu Met Pro His Tyr
700 705 710

20
tct acc ttt gcc tac ttc cca act gag gta aca cct cat gct ttt acc 2334
Ser Thr Phe Ala Tyr Phe Pro Thr Glu Val Thr Pro His Ala Phe Thr
715 720 725

25
cca tcc tcc aga caa cag gat ttg gtc tcc acg gtc aac gtg gta tac 2382
Pro Ser Ser Arg Gln Gln Asp Leu Val Ser Thr Val Asn Val Val Tyr
730 735 740 745

30
tcg cag aca acc caa ccg gta tac aat ggt gag aca cct ctt caa cct 2430
Ser Gln Thr Thr Gln Pro Val Tyr Asn Gly Glu Thr Pro Leu Gln Pro
750 755 760

35
tcc tac agt agt gaa gtc ttt cct cta gtc acc cct ttg ttg ctt gac 2478
Ser Tyr Ser Ser Glu Val Phe Pro Leu Val Thr Pro Leu Leu Leu Asp
765 770 775

40
aat cag atc ctc aac act acc cct gct gct tca agt agt gat tcg gcc 2526
Asn Gln Ile Leu Asn Thr Thr Pro Ala Ala Ser Ser Ser Asp Ser Ala
780 785 790

45
ttg cat gct acg cct gta ttt ccc agt gtc gat gtg tca ttt gaa tcc 2574
Leu His Ala Thr Pro Val Phe Pro Ser Val Asp Val Ser Phe Glu Ser
795 800 805

50
atc ctg tct tcc tat gat ggt gca cct ttg ctt cca ttt tcc tct gct 2622
Ile Leu Ser Ser Tyr Asp Gly Ala Pro Leu Leu Pro Phe Ser Ser Ala
810 815 820 825

55
tcc ttc agt agt gaa ttg ttt cgc cat ctg cat aca gtt tct caa atc 2670
Ser Phe Ser Ser Glu Leu Phe Arg His Leu His Thr Val Ser Gln Ile
830 835 840

60
ctt cca caa gtt act tca gct acc gag agt gat aag gtg ccc ttg cat 2718
Leu Pro Gln Val Thr Ser Ala Thr Glu Ser Asp Lys Val Pro Leu His
845 850 855

65
gct tct ctg cca gtg gct ggg ggt gat ttg cta tta gag ccc agc ctt 2766
Ala Ser Leu Pro Val Ala Gly Gly Asp Leu Leu Leu Glu Pro Ser Leu
860 865 870

70
gct cag tat tct gat gtg ctg tcc act act cat gct gct tca gag acg 2814
Ala Gln Tyr Ser Asp Val Leu Ser Thr Thr His Ala Ala Ser Glu Thr
875 880 885

5
10
15
20
25
30
35
40
45
50
55

ctg gaa ttt ggt agt gaa tct ggt gtt ctt tat aaa acg ctt atg ttt 2862
Leu Glu Phe Gly Ser Glu Ser Gly Val Leu Tyr Lys Thr Leu Met Phe
890 895 900 905

tct caa gtt gaa cca ccc agc agt gat gcc atg atg cat gca cgt tct 2910
Ser Gln Val Glu Pro Pro Ser Ser Asp Ala Met Met His Ala Arg Ser
910 915 920

tca ggg cct gaa cct tct tat gcc ttg tct gat aat gag ggc tcc caa 2958
Ser Gly Pro Glu Pro Ser Tyr Ala Leu Ser Asp Asn Glu Gly Ser Gln
925 930 935

cac atc ttc act gtt tct tac agt tct gca ata cct gtg cat gat tct 3006
His Ile Phe Thr Val Ser Tyr Ser Ser Ala Ile Pro Val His Asp Ser
940 945 950

gtg ggt gta act tat cag ggt tcc tta ttt agc ggc cct agc cat ata 3054
Val Gly Val Thr Tyr Gln Gly Ser Leu Phe Ser Gly Pro Ser His Ile
955 960 965

cca ata cct aag tct tcg tta ata acc cca act gca tca tta ctg cag 3102
Pro Ile Pro Lys Ser Ser Leu Ile Thr Pro Thr Ala Ser Leu Leu Gln
970 975 980 985

cct act cat gcc ctc tct ggt gat ggg gaa tgg tct gga gcc tct tct 3150
Pro Thr His Ala Leu Ser Gly Asp Gly Glu Trp Ser Ser Gly Ala Ser Ser
990 995 1000

gat agt gaa ttt ctt tta cct gac aca gat ggg ctg aca gcc ctt 3195
Asp Ser Glu Phe Leu Leu Pro Asp Thr Asp Gly Leu Thr Ala Leu
1005 1010 1015

aac att tct tca cct gtt tct gta gct gaa ttt aca tat aca aca 3240
Asn Ile Ser Ser Pro Val Ser Val Ala Glu Phe Thr Tyr Thr Thr
1020 1025 1030

tct gtg ttt ggt gat gat aat aag gcg ctt tct aaa agt gaa ata 3285
Ser Val Phe Gly Asp Asp Asn Lys Ala Leu Ser Lys Ser Glu Ile
1035 1040 1045

ata tat gga aat gag act gaa ctg caa att cct tct ttc aat gag 3330
Ile Tyr Gly Asn Glu Thr Glu Leu Gln Ile Pro Ser Phe Asn Glu
1050 1055 1060

atg gtt tac cct tct gaa agc aca gtc atg ccc aac atg tat gat 3375
Met Val Tyr Pro Ser Glu Ser Thr Val Met Pro Asn Met Tyr Asp
1065 1070 1075

aat gta aat aag ttg aat gcg tct tta caa gaa acc tct gtt tcc 3420
Asn Val Asn Lys Leu Asn Ala Ser Leu Gln Glu Thr Ser Val Ser
1080 1085 1090

att tct agc acc aag ggc atg ttt cca ggg tcc ctt gct cat acc 3465
Ile Ser Ser Thr Lys Gly Met Phe Pro Gly Ser Leu Ala His Thr
1095 1100 1105

5
10
15
20
25
30
35
40
45
50
55

acc act aag gtt ttt gat cat gag att agt caa gtt cca gaa aat	3510
Thr Thr Lys Val Phe Asp His Glu Ile Ser Gln Val Pro Glu Asn	
1110 1115 1120	
aac ttt tca gtt caa cct aca cat act gtc tct caa gca tct ggt	3555
Asn Phe Ser Val Gln Pro Thr His Thr Val Ser Gln Ala Ser Gly	
1125 1130 1135	
gac act tcg ctt aaa cct gtg ctt agt gca aac tca gag cca gca	3600
Asp Thr Ser Leu Lys Pro Val Leu Ser Ala Asn Ser Glu Pro Ala	
1140 1145 1150	
tcc tct gac cct gct tct agt gaa atg tta tct cct tca act cag	3645
Ser Ser Asp Pro Ala Ser Ser Glu Met Leu Ser Pro Ser Thr Gln	
1155 1160 1165	
ctc tta ttt tat gag acc tca gct tct ttt agt act gaa gta ttg	3690
Leu Leu Phe Tyr Glu Thr Ser Ala Ser Phe Ser Thr Glu Val Leu	
1170 1175 1180	
cta caa cct tcc ttt cag gct tct gat gtt gac acc ttg ctt aaa	3735
Leu Gln Pro Ser Phe Gln Ala Ser Asp Val Asp Thr Leu Leu Lys	
1185 1190 1195	
act gtt ctt cca gct gtg ccc agt gat cca ata ttg gtt gaa acc	3780
Thr Val Leu Pro Ala Val Pro Ser Asp Pro Ile Leu Val Glu Thr	
1200 1205 1210	
ccc aaa gtt gat aaa att agt tct aca atg ttg cat ctc att gta	3825
Pro Lys Val Asp Lys Ile Ser Ser Thr Met Leu His Leu Ile Val	
1215 1220 1225	
tca aat tct gct tca agt gaa aac atg ctg cac tct aca tct gta	3870
Ser Asn Ser Ala Ser Ser Glu Asn Met Leu His Ser Thr Ser Val	
1230 1235 1240	
cca gtt ttt gat gtg tcg cct act tct cat atg cac tct gct tca	3915
Pro Val Phe Asp Val Ser Pro Thr Ser His Met His Ser Ala Ser	
1245 1250 1255	
ctt caa ggt ttg acc att tcc tat gca agt gag aaa tat gaa cca	3960
Leu Gln Gly Leu Thr Ile Ser Tyr Ala Ser Glu Lys Tyr Glu Pro	
1260 1265 1270	
gtt ttg tta aaa agt gaa agt tcc cac caa gtg gta cct tct ttg	4005
Val Leu Leu Lys Ser Glu Ser Ser His Gln Val Val Pro Ser Leu	
1275 1280 1285	
tac agt aat gat gag ttg ttc caa acg gcc aat ttg gag att aac	4050
Tyr Ser Asn Asp Glu Leu Phe Gln Thr Ala Asn Leu Glu Ile Asn	
1290 1295 1300	
cag gcc cat ccc cca aaa gga agg cat gta ttt gct aca cct gtt	4095
Gln Ala His Pro Pro Lys Gly Arg His Val Phe Ala Thr Pro Val	
1305 1310 1315	

5 tta tca att gat gaa cca tta aat aca cta ata aat aag ctt ata 4140
Leu Ser Ile Asp Glu Pro Leu Asn Thr Leu Ile Asn Lys Leu Ile
1320 1325 1330

cat tcc gat gaa att tta acc tcc acc aaa agt tct gtt act ggt 4185
His Ser Asp Glu Ile Leu Thr Ser Thr Lys Ser Ser Val Thr Gly
1335 1340 1345

10 aag gta ttt gct ggt att cca aca gtt gct tct gat aca ttt gta 4230
Lys Val Phe Ala Gly Ile Pro Thr Val Ala Ser Asp Thr Phe Val
1350 1355 1360

15 tct act gat cat tct gtt cct ata gga aat ggg cat gtt gcc att 4275
Ser Thr Asp His Ser Val Pro Ile Gly Asn Gly His Val Ala Ile
1365 1370 1375

20 aca gct gtt tct ccc cac aga gat ggt tct gta acc tca aca aag 4320
Thr Ala Val Ser Pro His Arg Asp Gly Ser Val Thr Ser Thr Lys
1380 1385 1390

25 ttg ctg ttt cct tct aag gca act tct gag ctg agt cat agt gcc 4365
Leu Leu Phe Pro Ser Lys Ala Thr Ser Glu Leu Ser His Ser Ala
1395 1400 1405

30 aaa tct gat gcc ggt tta gtg ggt ggt ggt gaa gat ggt gac act 4410
Lys Ser Asp Ala Gly Leu Val Gly Gly Gly Glu Asp Gly Asp Thr
1410 1415 1420

gat gat gat ggt gat gat gat gat gac aga gat agt gat ggc tta 4455
Asp Asp Asp Gly Asp Asp Asp Asp Asp Arg Asp Ser Asp Gly Leu
1425 1430 1435

35 tcc att cat aag tgt atg tca tgc tca tcc tat aga gaa tca cag 4500
Ser Ile His Lys Cys Met Ser Cys Ser Ser Tyr Arg Glu Ser Gln
1440 1445 1450

40 gaa aag gta atg aat gat tca gac acc cac gaa aac agt ctt atg 4545
Glu Lys Val Met Asn Asp Ser Asp Thr His Glu Asn Ser Leu Met
1455 1460 1465

45 gat cag aat aat cca atc tca tac tca cta tct gag aat tct gaa 4590
Asp Gln Asn Asn Pro Ile Ser Tyr Ser Leu Ser Glu Asn Ser Glu
1470 1475 1480

gaa gat aat aga gtc aca agt gta tcc tca gac agt caa act ggt 4635
Glu Asp Asn Arg Val Thr Ser Val Ser Ser Asp Ser Gln Thr Gly
1485 1490 1495

50 atg gac aga agt cct ggt aaa tca cca tca gca aat ggg cta tcc 4680
Met Asp Arg Ser Pro Gly Lys Ser Pro Ser Ala Asn Gly Leu Ser
1500 1505 1510

55 caa aag cac aat gat gga aaa gag gaa aat gac att cag act ggt 4725
Gln Lys His Asn Asp Gly Lys Glu Glu Asn Asp Ile Gln Thr Gly
1515 1520 1525

5 agt gct ctg ctt cct ctc agc cct gaa tct aaa gca tgg gca gtt 4770
Ser Ala Leu Leu Pro Leu Ser Pro Glu Ser Lys Ala Trp Ala Val
1530 1535 1540

ctg aca agt gat gaa gaa agt gga tca ggg caa ggt acc tca gat 4815
Leu Thr Ser Asp Glu Glu Ser Gly Ser Gly Gln Gly Thr Ser Asp
1545 1550 1555

10 agc ctt aat gag aat gag act tcc aca gat ttc agt ttt gca gac 4860
Ser Leu Asn Glu Asn Glu Thr Ser Thr Asp Phe Ser Phe Ala Asp
1560 1565 1570

15 act aat gaa aaa gat gct gat ggg atc ctg gca gca ggt gac tca 4905
Thr Asn Glu Lys Asp Ala Asp Gly Ile Leu Ala Ala Gly Asp Ser
1575 1580 1585

20 gaa ata act cct gga ttc cca cag tcc cca aca tca tct gtt act 4950
Glu Ile Thr Pro Gly Phe Pro Gln Ser Pro Thr Ser Ser Val Thr
1590 1595 1600

25 agc gag aac tca gaa gtg ttc cac gtt tca gag gca gag gcc agt 4995
Ser Glu Asn Ser Glu Val Phe His Val Ser Glu Ala Glu Ala Ser
1605 1610 1615

30 aat agt agc cat gag tct cgt att ggt cta gct gag ggg ttg gaa 5040
Asn Ser Ser His Glu Ser Arg Ile Gly Leu Ala Glu Gly Leu Glu
1620 1625 1630

35 tcc gag aag aag gca gtt ata ccc ctt gtg atc gtg tca gcc ctg 5085
Ser Glu Lys Lys Ala Val Ile Pro Leu Val Ile Val Ser Ala Leu
1635 1640 1645

act ttt atc tgt cta gtg gtt ctt gtg ggt att ctc atc tac tgg 5130
Thr Phe Ile Cys Leu Val Val Leu Val Gly Ile Leu Ile Tyr Trp
1650 1655 1660

40 agg aaa tgc ttc cag act gca cac ttt tac tta gag gac agt aca 5175
Arg Lys Cys Phe Gln Thr Ala His Phe Tyr Leu Glu Asp Ser Thr
1665 1670 1675

tcc cct aga gtt ata tcc aca cct cca aca cct atc ttt cca att 5220
Ser Pro Arg Val Ile Ser Thr Pro Pro Thr Pro Ile Phe Pro Ile
1680 1685 1690

45 tca gat gat gtc gga gca att cca ata aag cac ttt cca aag cat 5265
Ser Asp Asp Val Gly Ala Ile Pro Ile Lys His Phe Pro Lys His
1695 1700 1705

50 gtt gca gat tta cat gca agt agt ggg ttt act gaa gaa ttt gag 5310
Val Ala Asp Leu His Ala Ser Ser Gly Phe Thr Glu Glu Phe Glu
1710 1715 1720

55 aca ctg aaa gag ttt tac cag gaa gtg cag agc tgt act gtt gac 5355
Thr Leu Lys Glu Phe Tyr Gln Glu Val Gln Ser Cys Thr Val Asp
1725 1730 1735

	tta ggt att aca gca gac agc tcc aac cac cca gac aac aag cac	5400
	Leu Gly Ile Thr 1740 Ala Asp Ser Ser Asn His Pro Asp Asn Lys His 1750	
5	aag aat cga tac ata aat atc gtt gcc tat gat cat agc agg gtt	5445
	Lys Asn Arg Tyr 1755 Ile Asn Ile Val Ala Tyr Asp His Ser Arg Val 1765	
10	aag cta gca cag ctt gct gaa aag gat ggc aaa ctg act gat tat	5490
	Lys Leu Ala Gln 1770 Leu Ala Glu Lys Asp Gly Lys Leu Thr Asp Tyr 1780	
15	atc aat gcc aat tat gtt gat ggc tac aac aga cca aaa gct tat	5535
	Ile Asn Ala Asn 1785 Tyr Val Asp Gly Tyr Asn Arg Pro Lys Ala Tyr 1795	
20	att gct gcc caa ggc cca ctg aaa tcc aca gct gaa gat ttc tgg	5580
	Ile Ala Ala Gln 1800 Gly Pro Leu Lys Ser Thr Ala Glu Asp Phe Trp 1810	
25	aga atg ata tgg gaa cat aat gtg gaa gtt att gtc atg ata aca	5625
	Arg Met Ile Trp 1815 Glu His Asn Val Glu Val Ile Val Met Ile Thr 1825	
30	aac ctc gtg gag aaa gga agg aga aaa tgt gat cag tac tgg cct	5670
	Asn Leu Val Glu 1830 Lys Gly Arg Arg Lys Cys Asp Gln Tyr Trp Pro 1840	
35	gcc gat ggg agt gag gag tac ggg aac ttt ctg gtc act cag aag	5715
	Ala Asp Gly Ser 1845 Glu Glu Tyr Gly Asn Phe Leu Val Thr Gln Lys 1855	
40	agt gtg caa gtg ctt gcc tat tat act gtg agg aat ttt act cta	5760
	Ser Val Gln Val 1860 Leu Ala Tyr Tyr Thr Val Arg Asn Phe Thr Leu 1870	
45	aga aac aca aaa ata aaa aag ggc tcc cag aaa gga aga ccc agt	5805
	Arg Asn Thr Lys 1875 Ile Lys Lys Gly Ser Gln Lys Gly Arg Pro Ser 1885	
50	gga cgt gtg gtc aca cag tat cac tac acg cag tgg cct gac atg	5850
	Gly Arg Val Val 1890 Thr Gln Tyr His Tyr Thr Gln Trp Pro Asp Met 1900	
55	gga gta cca gag tac tcc ctg cca gtg ctg acc ttt gtg aga aag	5895
	Gly Val Pro Glu 1905 Tyr Ser Leu Pro Val Leu Thr Phe Val Arg Lys 1915	
60	gca gcc tat gcc aag cgc cat gca gtg ggg cct gtt gtc gtc cac	5940
	Ala Ala Tyr Ala 1920 Lys Arg His Ala Val Gly Pro Val Val Val His 1930	
65	tgc agt gct gga gtt gga aga aca ggc aca tat att gtg cta gac	5985
	Cys Ser Ala Gly 1935 Val Gly Arg Thr Gly Thr Tyr Ile Val Leu Asp 1945	

5 agt atg ttg cag cag att caa cac gaa gga act gtc aac ata ttt 6030
Ser Met Leu Gln Gln Ile Gln His Glu Gly Thr Val Asn Ile Phe
1950 1955 1960

ggc ttc tta aaa cac atc cgt tca caa aga aat tat ttg gta caa 6075
Gly Phe Leu Lys His Ile Arg Ser Gln Arg Asn Tyr Leu Val Gln
1965 1970 1975

10 act gag gag caa tat gtc ttc att cat gat aca ctg gtt gag gcc 6120
Thr Glu Glu Gln Tyr Val Phe Ile His Asp Thr Leu Val Glu Ala
1980 1985 1990

15 ata ctt agt aaa gaa act gag gtg ctg gac agt cat att cat gcc 6165
Ile Leu Ser Lys Glu Thr Glu Val Leu Asp Ser His Ile His Ala
1995 2000 2005

20 tat gtt aat gca ctc ctc att cct gga cca gca ggc aaa aca aag 6210
Tyr Val Asn Ala Leu Leu Ile Pro Gly Pro Ala Gly Lys Thr Lys
2010 2015 2020

25 cta gag aaa caa ttc cag ctc ctg agc cag tca aat ata cag cag 6255
Leu Glu Lys Gln Phe Gln Leu Leu Ser Gln Ser Asn Ile Gln Gln
2025 2030 2035

30 agt gac tat tct gca gcc cta aag caa tgc aac agg gaa aag aat 6300
Ser Asp Tyr Ser Ala Ala Leu Lys Gln Cys Asn Arg Glu Lys Asn
2040 2045 2050

35 cga act tct tct atc atc cct gtg gaa aga tca agg gtt ggc att 6345
Arg Thr Ser Ser Ile Ile Pro Val Glu Arg Ser Arg Val Gly Ile
2055 2060 2065

40 tca tcc ctg agt gga gaa ggc aca gac tac atc aat gcc tcc tat 6390
Ser Ser Leu Ser Gly Glu Gly Thr Asp Tyr Ile Asn Ala Ser Tyr
2070 2075 2080

atc atg ggc tat tac cag agc aat gaa ttc atc att acc cag cac 6435
Ile Met Gly Tyr Tyr Gln Ser Asn Glu Phe Ile Ile Thr Gln His
2085 2090 2095

45 cct ctc ctt cat acc atc aag gat ttc tgg agg atg ata tgg gac 6480
Pro Leu Leu His Thr Ile Lys Asp Phe Trp Arg Met Ile Trp Asp
2100 2105 2110

cat aat gcc caa ctg gtg gtt atg att cct gat ggc caa aac atg 6525
His Asn Ala Gln Leu Val Val Met Ile Pro Asp Gly Gln Asn Met
2115 2120 2125

50 gca gaa gat gaa ttt gtt tac tgg cca aat aaa gat gag cct ata 6570
Ala Glu Asp Glu Phe Val Tyr Trp Pro Asn Lys Asp Glu Pro Ile
2130 2135 2140

55 aat tgt gag agc ttt aag gtc act ctt atg gct gaa gaa cac aaa 6615
Asn Cys Glu Ser Phe Lys Val Thr Leu Met Ala Glu Glu His Lys
2145 2150 2155

5 tgt cta tct aat gag gaa aaa ctt ata att cag gac ttt atc tta 6660
Cys Leu Ser Asn Glu Glu Lys Leu Ile Ile Gln Asp Phe Ile Leu
2160 2165 2170

gaa gct aca cag gat gat tat gta ctt gaa gtg agg cac ttt cag 6705
Glu Ala Thr Gln Asp Asp Tyr Val Leu Glu Val Arg His Phe Gln
2175 2180 2185

10 tgt cct aaa tgg cca aat cca gat agc ccc att agt aaa act ttt 6750
Cys Pro Lys Trp Pro Asn Pro Asp Ser Pro Ile Ser Lys Thr Phe
2190 2195 2200

15 gaa ctt ata agt gtt ata aaa gaa gaa gct gcc aat agg gat ggg 6795
Glu Leu Ile Ser Val Ile Lys Glu Glu Ala Ala Asn Arg Asp Gly
2205 2210 2215

20 cct atg att gtt cat gat gag cat gga gga gtg acg gca gga act 6840
Pro Met Ile Val His Asp Glu His Gly Gly Val Thr Ala Gly Thr
2220 2225 2230

25 ttc tgt gct ctg aca acc ctt atg cac caa cta gaa aaa gaa aat 6885
Phe Cys Ala Leu Thr Thr Leu Met His Gln Leu Glu Lys Glu Asn
2235 2240 2245

tcc gtg gat gtt tac cag gta gcc aag atg atc aat ctg atg agg 6930
Ser Val Asp Val Tyr Gln Val Ala Lys Met Ile Asn Leu Met Arg
2250 2255 2260

30 cca gga gtc ttt gct gac att gag cag tat cag ttt ctc tac aaa 6975
Pro Gly Val Phe Ala Asp Ile Glu Gln Tyr Gln Phe Leu Tyr Lys
2265 2270 2275

35 gtg atc ctc agc ctt gtg agc aca agg cag gaa gag aat cca tcc 7020
Val Ile Leu Ser Leu Val Ser Thr Arg Gln Glu Glu Asn Pro Ser
2280 2285 2290

40 acc tct ctg gac agt aat ggt gca gca ttg cct gat gga aat ata 7065
Thr Ser Leu Asp Ser Asn Gly Ala Ala Leu Pro Asp Gly Asn Ile
2295 2300 2305

gct gag agc tta gag tct tta gtt taa cacagaaagg ggtgggggga 7112
Ala Glu Ser Leu Glu Ser Leu Val
2310

45 ctacacatctg agcattgttt tcctcttcct aaaattaggc aggaaaaatca gtctagttct 7172

gttatctgtt gatttcccat cacctgacag taactttcat gacataggat tctgccgcca 7232

50 aatttatatc attaacaatg tgtgcctttt tgcaagactt gtaatttact tattatgttt 7292

gaactaaaaat gattgaattt tacagtattt ctaagaatgg aattgtggta tttttttctg 7352

55 tattgatttt aacagaaaaat ttcaatttat agaggttagg aattccaaac tacagaaaaat 7412

gtttgttttt agtgtcaaat ttttagctgt attttagtagca attatcaggt ttgctagaaa 7472

tataactttt aatacagtag cctgtaaata aaacactctt ccatatgata ttcaacattt 7532
 5 tacaactgca gtattcacct aaagtagaaa taatctgtta cttattgtaa atactgccct 7592
 agtgtctcca tggaccaaatt ttatatattat aattgtagat ttttatattt tactactgag 7652
 tcaagttttc tagttctgtg taattgttta gtttaatgac gtagttcatt agctggctctt 7712
 10 actctaccag ttttctgaca ttgtattgtg ttacctaatg cattaacttt gtttcagcat 7772
 gtaattttta cttttgtgga aaatagaaat accttcattt tgaaagaagt ttttatgaga 7832
 ataacacctt accaaacatt gttcaaattg tttttatcca aggaattgca aaaataaata 7892
 15 taaatattgc cattaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 7941

<210> 6

<211> 2314

<212> PRT

<213> Homo sapiens

<220>

<221> Reference

<222> (1)..(2314)

35 <223> Krueger, N.X. and Saito, H.: A human transmembrane protein-tyrosi
 ne-phosphatase, PTP zeta, is expressed in brain and has an N-term
 inal receptor domain homologous to carbonic anhydrases; Proc. Nat
 l. Acad. Sci. USA 89 (16), 7417-7421 (1992)

<220>

<221> gene

<222> (1)..(2314)

<223> PTP-zeta

<220>

<221> Reference

<222> (1)..(2314)

<223> Levy, J.B., et al.; The cloning of a receptor-type protein tyrosine phosphatase expressed in the central nervous system; J. Biol. Chem. 268 (14), 10573-10581 (1993)

5

<220>

10 <221> SIGNAL

<222> (1)..(24)

15

<223> By similarity

<220>

20 <221> DOMAIN

<222> (25)..(1635)

25

<223> Extracellular (potential)

<220>

30 <221> MISC_FEATURE

<222> (25)..(2314)

35

<223> Mature chain; protein-tyrosine phosphatase zeta

<220>

40 <221> DOMAIN

<222> (34)..(302)

45

<223> Carbonic-anhydrase like

<220>

50 <221> SITE

<222> (105)..(105)

55

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (134)..(134)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (223)..(223)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (232)..(232)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (324)..(324)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> DOMAIN

<222> (312)..(406)

<223> Fibronectin Type-III

<220>

<221> SITE

<222> (381)..(381)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

5

<220>

<221> SITE

10

<222> (497)..(497)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

15

<220>

<221> SITE

20

<222> (501)..(501)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

25

<220>

<221> SITE

30

<222> (552)..(552)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

35

<220>

<221> BINDING

40

<222> (587)..(587)

<223> Chondrrroitin Sulfate (potential)

45

<220>

<221> SITE

50

<222> (602)..(602)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

55

<220>

5
10
15
20
25
30
35
40
45
50
55

<221> SITE
<222> (629)..(629)
<223> Glycosylation; N-linked (GLCNAC...) (potential)
<220>
<221> BINDING
<222> (637)..(637)
<223> Chondroitin sulfate (potential)
<220>
<221> SITE
<222> (677)..(677)
<223> Glycosylation; N-linked (GLCNAC...) (potential)
<220>
<221> VARSPLIC
<222> (755)..(1614)
<223> Splicing variant; missing (in short isoform)
<220>
<221> BINDING
<222> (997)..(997)
<223> Chondroitin sulfate (potential)
<220>
<221> SITE
<222> (1017)..(1017)
<223> Glycosylation; N-linked (GLCNAC...) (potential)

5 <220>
<221> SITE
<222> (1050)..(1050)
10 <223> Glycosylation; N-linked (GLCNAC...) (potential)

15 <220>
<221> SITE
<222> (1082)..(1082)
20 <223> Glycosylation; N-linked (GLCNAC...) (potential)

25 <220>
<221> SITE
<222> (1122)..(1122)
30 <223> Glycosylation; N-linked (GLCNAC...) (potential)

35 <220>
<221> SITE
<222> (1456)..(1456)
40 <223> Glycosylation; N-linked (GLCNAC...) (potential)

45 <220>
<221> BINDING
<222> (1548)..(1548)
50 <223> Chondroitin sulfate (potential)

55 <220>
<221> BINDING

<222> (1550)..(1550)

<223> Chondroitin sulfate (potential)

<220>

<221> SITE

<222> (1561)..(1561)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (1617)..(1617)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> TRANSMEM

<222> (1636)..(1661)

<223> Transmembrane region; potential

<220>

<221> DOMAIN

<222> (1662)..(2314)

<223> Cytoplasmic (potential)

<220>

<221> CONFLICT

<222> (1722)..(1728)

<223> Missing (in ref. 2)

<220>

<221> DOMAIN

<222> (1744)..(1997)

<223> Protein-tyrosine phosphates

<220>

<221> ACT_SITE

<222> (1932)..(1932)

<223> Active site; by similarity

<220>

<221> DOMAIN

<222> (1998)..(2314)

<223> Protein-tyrosine phosphatase

<220>

<221> MISC_FEATURE

<222> (2222)..(2222)

<223> Ancestral active site

<400> 6

Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys Val
1 5 10 15

Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg Lys Leu
20 25 30

Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn Gln Lys Asn
35 40 45

Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys Gln Ser Pro Ile

50

55

60

5 Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu
65 70 75 80

10

Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His
85 90 95

15

Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val
100 105 110

Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe
115 120 125

20 His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu
130 135 140

25 Glu Gly Gln Lys Phe Pro Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala
145 150 155 160

30 Asp Arg Phe Ser Ser Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu
165 170 175

Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp
180 185 190

35 Phe Lys Ala Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys
195 200 205

40 Gln Ala Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser
210 215 220

45 Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
225 230 235 240

50 Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser
245 250 255

Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser
260 265 270

55 Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln

275

280

285

5 Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu
290 295 300

10 Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala
305 310 315 320

15 Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg
325 330 335
Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln
340 345 350

20 Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr
355 360 365

25 Gln Asp Leu Gly Ala Ile Leu Asn Asn Leu Leu Pro Asn Met Ser Tyr
370 375 380

30 Val Leu Gln Ile Val Ala Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr
385 390 395 400

Ser Asp Gln Leu Ile Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp
405 410 415

35 Leu Phe Pro Glu Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu
420 425 430

40 Glu Gly Lys Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp
435 440 445

45 Ser Ala Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr
450 455 460

50 Thr His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn
465 470 475 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val Pro
485 490 495

55 Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu Ala Thr

500

505

510

5 Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu Leu Pro Pro
515 520 525

10 His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp Gly Ser Lys Thr
530 535 540

Val Leu Arg Ser Pro His Met Asn Leu Ser Gly Thr Ala Glu Ser Leu
545 550 555 560

15 Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu Glu Ser Leu Leu Thr Ser
565 570 575

20 Phe Lys Leu Asp Thr Gly Ala Glu Asp Ser Ser Gly Ser Ser Pro Ala
580 585 590

25 Thr Ser Ala Ile Pro Phe Ile Ser Glu Asn Ile Ser Gln Gly Tyr Ile
595 600 605

30 Phe Ser Ser Glu Asn Pro Glu Thr Ile Thr Tyr Asp Val Leu Ile Pro
610 615 620

Glu Ser Ala Arg Asn Ala Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu
625 630 635 640

35 Glu Ser Leu Lys Asp Pro Ser Met Glu Gly Asn Val Trp Phe Pro Ser
645 650 655

40 Ser Thr Asp Ile Thr Ala Gln Pro Asp Val Gly Ser Gly Arg Glu Ser
660 665 670

45 Phe Leu Gln Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys
675 680 685

50 Thr Thr Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser
690 695 700

Val Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro
705 710 715 720

55 Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln Asp

725

730

735

5

Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln Pro Val
740 745 750

10

Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser Glu Val Phe
755 760 765

15

Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile Leu Asn Thr Thr
770 775 780

Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His Ala Thr Pro Val Phe
785 790 795 800

20

Pro Ser Val Asp Val Ser Phe Glu Ser Ile Leu Ser Ser Tyr Asp Gly
805 810 815

25

Ala Pro Leu Leu Pro Phe Ser Ser Ala Ser Phe Ser Ser Glu Leu Phe
820 825 830

30

Arg His Leu His Thr Val Ser Gln Ile Leu Pro Gln Val Thr Ser Ala
835 840 845

Thr Glu Ser Asp Lys Val Pro Leu His Ala Ser Leu Pro Val Ala Gly
850 855 860

35

Gly Asp Leu Leu Leu Glu Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu
865 870 875 880

40

Ser Thr Thr His Ala Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser
885 890 895

45

Gly Val Leu Tyr Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser
900 905 910

50

Ser Asp Ala Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr
915 920 925

Ala Leu Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr
930 935 940

55

Ser Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly

960

Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr Glu Thr Ser

1160 1165 1170

5 Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser Phe Gln Ala
1175 1180 1185

10 Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro Ala Val Pro
1190 1195 1200

15 Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp Lys Ile Ser
1205 1210 1215

20 Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala Ser Ser Glu
1220 1225 1230

25 Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp Val Ser Pro
1235 1240 1245

30 Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu Thr Ile Ser
1250 1255 1260

35 Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys Ser Glu Ser
1265 1270 1275

40 Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp Glu Leu Phe
1280 1285 1290

45 Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro Pro Lys Gly
1295 1300 1305

50 Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp Glu Pro Leu
1310 1315 1320

55 Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu Ile Leu Thr
1325 1330 1335

Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala Gly Ile Pro
1340 1345 1350

Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His Ser Val Pro
1355 1360 1365

Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser Pro His Arg

1370 1375 1380

5 Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro Ser Lys Ala
1385 1390 1395

10 Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala Gly Leu Val
1400 1405 1410

15 Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly Asp Asp Asp
1415 1420 1425

20 Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys Cys Met Ser
1430 1435 1440

25 Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met Asn Asp Ser
1445 1450 1455

30 Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn Pro Ile Ser
1460 1465 1470

35 Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg Val Thr Ser
1475 1480 1485

40 Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser Pro Gly Lys
1490 1495 1500

45 Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn Asp Gly Lys
1505 1510 1515

50 Glu Glu Asn Asp Ile Gln Thr Gly Ser Ala Leu Leu Pro Leu Ser
1520 1525 1530

55 Pro Glu Ser Lys Ala Trp Ala Val Leu Thr Ser Asp Glu Glu Ser
1535 1540 1545

60 Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu Asn Glu Thr
1550 1555 1560

65 Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys Asp Ala Asp
1565 1570 1575

70 Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro Gly Phe Pro

1580 1585 1590

5 Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser Glu Val Phe
1595 1600 1605

10 His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His Glu Ser Arg
1610 1615 1620

Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys Ala Val Ile
1625 1630 1635

15 Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys Leu Val Val
1640 1645 1650

20 Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe Gln Thr Ala
1655 1660 1665

25 His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val Ile Ser Thr
1670 1675 1680

30 Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val Gly Ala Ile
1685 1690 1695

35 Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu His Ala Ser
1700 1705 1710

Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu Phe Tyr Gln
1715 1720 1725

40 Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr Ala Asp Ser
1730 1735 1740

45 Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr Ile Asn Ile
1745 1750 1755

50 Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln Leu Ala Glu
1760 1765 1770

Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn Tyr Val Asp
1775 1780 1785

55 Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln Gly Pro Leu

1790 1795 1800

5 Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp Glu His Asn
1805 1810 1815

10 Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu Lys Gly Arg
1820 1825 1830

15 Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser Glu Glu Tyr
1835 1840 1845

Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val Leu Ala Tyr
1850 1855 1860

20 Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys Ile Lys Lys
1865 1870 1875

25 Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val Thr Gln Tyr
1880 1885 1890

30 His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu Tyr Ser Leu
1895 1900 1905

35 Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala Lys Arg His
1910 1915 1920

Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly Val Gly Arg
1925 1930 1935

40 Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln Gln Ile Gln
1940 1945 1950

45 His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys His Ile Arg
1955 1960 1965

50 Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln Tyr Val Phe
1970 1975 1980

Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys Glu Thr Glu
1985 1990 1995

55 Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala Leu Leu Ile

2000 2005 2010

5 Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln Phe Gln Leu
2015 2020 2025

10 Leu Ser Gln Ser Asn Ile Gln Gln Ser Asp Tyr Ser Ala Ala Leu
2030 2035 2040

15 Lys Gln Cys Asn Arg Glu Lys Asn Arg Thr Ser Ser Ile Ile Pro
2045 2050 2055

20 Val Glu Arg Ser Arg Val Gly Ile Ser Ser Leu Ser Gly Glu Gly
2060 2065 2070

25 Thr Asp Tyr Ile Asn Ala Ser Tyr Ile Met Gly Tyr Tyr Gln Ser
2075 2080 2085

30 Asn Glu Phe Ile Ile Thr Gln His Pro Leu Leu His Thr Ile Lys
2090 2095 2100

35 Asp Phe Trp Arg Met Ile Trp Asp His Asn Ala Gln Leu Val Val
2105 2110 2115

40 Met Ile Pro Asp Gly Gln Asn Met Ala Glu Asp Glu Phe Val Tyr
2120 2125 2130

45 Trp Pro Asn Lys Asp Glu Pro Ile Asn Cys Glu Ser Phe Lys Val
2135 2140 2145

50 Thr Leu Met Ala Glu Glu His Lys Cys Leu Ser Asn Glu Glu Lys
2150 2155 2160

55 Leu Ile Ile Gln Asp Phe Ile Leu Glu Ala Thr Gln Asp Asp Tyr
2165 2170 2175

Val Leu Glu Val Arg His Phe Gln Cys Pro Lys Trp Pro Asn Pro
2180 2185 2190

Asp Ser Pro Ile Ser Lys Thr Phe Glu Leu Ile Ser Val Ile Lys
2195 2200 2205

Glu Glu Ala Ala Asn Arg Asp Gly Pro Met Ile Val His Asp Glu

2210

2215

2220

5 His Gly Gly Val Thr Ala Gly Thr Phe Cys Ala Leu Thr Thr Leu
2225 2230 2235

10 Met His Gln Leu Glu Lys Glu Asn Ser Val Asp Val Tyr Gln Val
2240 2245 2250

Ala Lys Met Ile Asn Leu Met Arg Pro Gly Val Phe Ala Asp Ile
2255 2260 2265

15 Glu Gln Tyr Gln Phe Leu Tyr Lys Val Ile Leu Ser Leu Val Ser
2270 2275 2280

20 Thr Arg Gln Glu Glu Asn Pro Ser Thr Ser Leu Asp Ser Asn Gly
2285 2290 2295

25 Ala Ala Leu Pro Asp Gly Asn Ile Ala Glu Ser Leu Glu Ser Leu
2300 2305 2310

Val

<210> 7

<211> 1518

<212> DNA

<213> Homo sapiens

<220>

<221> gene

<222> (1)..(1518)

<223> Angiopoietin-like 2 (ANGPTL2), mRNA

<220>

<221> Reference

<222> (1)..(1518)

<223> Kim, I., Moon, S.O., Koh, K.N., Kim, H., Uhm, C.S., Kwak, H.J., Kim, N.G. and Koh, G.Y., "Molecular cloning, expression, and characterization of angiopoietin-related protein. angiopoietin-related protein induces endothelial cell sprouting", J. Biol. Chem. 274 (37), 26523-26528 (1999)

<220>

<221> CDS

<222> (22)..(1503)

<223>

<400> 7

aaccaccatt ttgcaaggac c atg agg cca ctg tgc gtg aca tgc tgg tgg 51
Met Arg Pro Leu Cys Val Thr Cys Trp Trp
1 5 10

ctc gga ctg ctg gct gcc atg gga gct gtt gca ggc cag gag gac ggt 99
Leu Gly Leu Leu Ala Ala Met Gly Ala Val Ala Gly Gln Glu Asp Gly
15 20 25

ttt gag ggc act gag gag ggc tcg cca aga gag ttc att tac cta aac 147
Phe Glu Gly Thr Glu Glu Gly Ser Pro Arg Glu Phe Ile Tyr Leu Asn
30 35 40

agg tac aag cgg gcg ggc gag tcc cag gac aag tgc acc tac acc ttc 195
Arg Tyr Lys Arg Ala Gly Glu Ser Gln Asp Lys Cys Thr Tyr Thr Phe
45 50 55

att gtg ccc cag cag cgg gtc acg ggt gcc atc tgc gtc aac tcc aag 243
Ile Val Pro Gln Gln Arg Val Thr Gly Ala Ile Cys Val Asn Ser Lys
60 65 70

gag cct gag gtg ctt ctg gag aac cga gtg cat aag cag gag cta gag 291
Glu Pro Glu Val Leu Leu Glu Asn Arg Val His Lys Gln Glu Leu Glu
75 80 85 90

ctg ctc aac aat gag ctg ctc aag cag aag cgg cag atc gag aca ctg 339
Leu Leu Asn Asn Glu Leu Leu Lys Gln Lys Arg Gln Ile Glu Thr Leu
95 100 105

cag cag ctg gtg gag gtg gac ggc ggc att gtg agc gag gtg aag ctg 387
Gln Gln Leu Val Glu Val Asp Gly Gly Ile Val Ser Glu Val Lys Leu
110 115 120

ctg cgc aag gag agc cgc aac atg aac tcg cgg gtc acg cag ctc tac 435
Leu Arg Lys Glu Ser Arg Asn Met Asn Ser Arg Val Thr Gln Leu Tyr
125 130 135

atg cag ctc ctg cac gag atc atc cgc aag cgg gac aac gcg ttg gag 483
Met Gln Leu Leu His Glu Ile Ile Arg Lys Arg Asp Asn Ala Leu Glu
140 145 150

5 ctc tcc cag ctg gag aac agg atc ctg aac cag aca gcc gac atg ctg 531
Leu Ser Gln Leu Glu Asn Arg Ile Leu Asn Gln Thr Ala Asp Met Leu
155 160 165 170

10 cag ctg gcc agc aag tac aag gac ctg gag cac aag tac cag cac ctg 579
Gln Leu Ala Ser Lys Tyr Lys Asp Leu Glu His Lys Tyr Gln His Leu
175 180 185

15 gcc aca ctg gcc cac aac caa tca gag atc atc gcg cag ctt gag gag 627
Ala Thr Leu Ala His Asn Gln Ser Glu Ile Ile Ala Gln Leu Glu Glu
190 195 200

20 cac tgc cag agg gtg ccc tcg gcc agg ccc gtc ccc cag cca ccc ccc 675
His Cys Gln Arg Val Pro Ser Ala Arg Pro Val Pro Gln Pro Pro Pro
205 210 215

25 gct gcc ccg ccc cgg gtc tac caa cca ccc acc tac aac cgc atc atc 723
Ala Ala Pro Pro Arg Val Tyr Gln Pro Pro Thr Tyr Asn Arg Ile Ile
220 225 230

30 aac cag atc tct acc aac gag atc cag agt gac cag aac ctg aag gtg 771
Asn Gln Ile Ser Thr Asn Glu Ile Gln Ser Asp Gln Asn Leu Lys Val
235 240 245 250

35 ctg cca ccc cct ctg ccc act atg ccc act ctc acc agc ctc cca tct 819
Leu Pro Pro Pro Leu Pro Thr Met Pro Thr Leu Thr Ser Leu Pro Ser
255 260 265

40 tcc acc gac aag ccg tcg ggc cca tgg aga gac tgc ctg cag gcc ctg 867
Ser Thr Asp Lys Pro Ser Gly Pro Trp Arg Asp Cys Leu Gln Ala Leu
270 275 280

45 gag gat ggc cac gac acc agc tcc atc tac ctg gtg aag ccg gag aac 915
Glu Asp Gly His Asp Thr Ser Ser Ile Tyr Leu Val Lys Pro Glu Asn
285 290 295

50 acc aac cgc ctc atg cag gtg tgg tgc gac cag aga cac gac ccc ggg 963
Thr Asn Arg Leu Met Gln Val Trp Cys Asp Gln Arg His Asp Pro Gly
300 305 310

55 ggc tgg acc gtc atc cag aga cgc ctg gat ggc tct gtt aac ttc ttc 1011
Gly Trp Thr Val Ile Gln Arg Arg Leu Asp Gly Ser Val Asn Phe Phe
315 320 325 330

agg aac tgg gag acg tac aag caa ggg ttt ggg aac att gat ggc gaa 1059
Arg Asn Trp Glu Thr Tyr Lys Gln Gly Phe Gly Asn Ile Asp Gly Glu
335 340 345

55 tac tgg ctg ggc ctg gag aac att tac tgg ctg acg aac caa ggc aac 1107
Tyr Trp Leu Gly Leu Glu Asn Ile Tyr Trp Leu Thr Asn Gln Gly Asn
350 355 360

5 gca gaa tac gcc agt ttc cgc ctg gaa cct gag agc gag tat tat aag 1203
Ala Glu Tyr Ala Ser Phe Arg Leu Glu Pro Glu Ser Glu Tyr Tyr Lys
380 385 390

15 cac aac ggc aag cag ttc acc acc ctg gac aga gat cat gat gtc tac 1299
His Asn Gly Lys Gln Phe Thr Thr Leu Asp Arg Asp His Asp Val Tyr
 415 420 425

tgt gcc cac tcc aac ctc aac ggg gtc tgg tac cgc ggg ggc cat tac 1395
 Cys Ala His Ser Asn Leu Asn Gly Val Trp Tyr Arg Gly Gly His Tyr
 445 450 455

30 tct tac tca ctc aag aaa gtg gtg atg atg atc cga ccg aac ccc aac 1491
Ser Tyr Ser Leu Lys Lys Val Val Met Met Ile Arg Pro Asn Pro Asn
475 480 485 490

<210> 8

<212> PRT

<220>

<222> (1) .. (493)

OC-94692.1

<220>

<221> SIGNAL

<222> (1)..(22)

<223> Potential

<220>

<221> CHAIN

<222> (23)..(493)

<223> Angiopoietin-related protein 2

<220>

<221> DOMAIN

<222> (76)..(115)

<223> Coiled Coil (potential)

<220>

<221> DOMAIN

<222> (152)..(206)

<223> Coiled Coil (potential)

<220>

<221> DOMAIN

<222> (438)..(450)

<223> Fibrinogen C-terminal

<220>

<221> CARBOHYD

<222> (164)..(164)

<223> N-linked (GLCNAC...) (potential)

5

<220>

<221> CARBOHYD

10

<222> (192)..(192)

<223> N-linked (GLCNAC...) (potential)

15

<400> 8

Met Arg Pro Leu Cys Val Thr Cys Trp Trp Leu Gly Leu Leu Ala Ala
1 5 10 15

Met Gly Ala Val Ala Gly Gln Glu Asp Gly Phe Glu Gly Thr Glu Glu
20 25 30

Gly Ser Pro Arg Glu Phe Ile Tyr Leu Asn Arg Tyr Lys Arg Ala Gly
35 40 45

Glu Ser Gln Asp Lys Cys Thr Tyr Thr Phe Ile Val Pro Gln Gln Arg
50 55 60

Val Thr Gly Ala Ile Cys Val Asn Ser Lys Glu Pro Glu Val Leu Leu
65 70 75 80

Glu Asn Arg Val His Lys Gln Glu Leu Glu Leu Leu Asn Asn Glu Leu
85 90 95

Leu Lys Gln Lys Arg Gln Ile Glu Thr Leu Gln Gln Leu Val Glu Val
100 105 110

Asp Gly Gly Ile Val Ser Glu Val Lys Leu Leu Arg Lys Glu Ser Arg
115 120 125

Asn Met Asn Ser Arg Val Thr Gln Leu Tyr Met Gln Leu Leu His Glu
130 135 140

Ile Ile Arg Lys Arg Asp Asn Ala Leu Glu Leu Ser Gln Leu Glu Asn
145 150 155 160

Arg Ile Leu Asn Gln Thr Ala Asp Met Leu Gln Leu Ala Ser Lys Tyr
165 170 175

5 Lys Asp Leu Glu His Lys Tyr Gln His Leu Ala Thr Leu Ala His Asn
180 185 190

10 Gln Ser Glu Ile Ile Ala Gln Leu Glu Glu His Cys Gln Arg Val Pro
195 200 205

15 Ser Ala Arg Pro Val Pro Gln Pro Pro Pro Ala Ala Pro Pro Arg Val
210 215 220

20 Tyr Gln Pro Pro Thr Tyr Asn Arg Ile Ile Asn Gln Ile Ser Thr Asn
225 230 235 240

25 Glu Ile Gln Ser Asp Gln Asn Leu Lys Val Leu Pro Pro Pro Leu Pro
245 250 255

30 Thr Met Pro Thr Leu Thr Ser Leu Pro Ser Ser Thr Asp Lys Pro Ser
260 265 270

35 Gly Pro Trp Arg Asp Cys Leu Gln Ala Leu Glu Asp Gly His Asp Thr
275 280 285

40 Ser Ser Ile Tyr Leu Val Lys Pro Glu Asn Thr Asn Arg Leu Met Gln
290 295 300

45 Val Trp Cys Asp Gln Arg His Asp Pro Gly Gly Trp Thr Val Ile Gln
305 310 315 320

50 Arg Arg Leu Asp Gly Ser Val Asn Phe Phe Arg Asn Trp Glu Thr Tyr
325 330 335

55 Lys Gln Gly Phe Gly Asn Ile Asp Gly Glu Tyr Trp Leu Gly Leu Glu
340 345 350

60 Asn Ile Tyr Trp Leu Thr Asn Gln Gly Asn Tyr Lys Leu Leu Val Thr
355 360 365

65 Met Glu Asp Trp Ser Gly Arg Lys Val Phe Ala Glu Tyr Ala Ser Phe
370 375 380

Arg Leu Glu Pro Glu Ser Glu Tyr Tyr Lys Leu Arg Leu Gly Arg Tyr
385 390 395 400

5

His Gly Asn Ala Gly Asp Ser Phe Thr Trp His Asn Gly Lys Gln Phe
405 410 415

10

Thr Thr Leu Asp Arg Asp His Asp Val Tyr Thr Gly Asn Cys Ala His
420 425 430

15

Tyr Gln Lys Gly Gly Trp Trp Tyr Asn Ala Cys Ala His Ser Asn Leu
435 440 445

20

Asn Gly Val Trp Tyr Arg Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp
450 455 460

25

Gly Val Tyr Trp Ala Glu Phe Arg Gly Gly Ser Tyr Ser Leu Lys Lys
465 470 475 480

30

Val Val Met Met Ile Arg Pro Asn Pro Asn Thr Phe His
485 490

35

<210> 9

<211> 2133

<212> DNA

<213> Homo sapiens

40

<220>

<221> gene

<222> (1) .. (2133)

45

<223> Human SPARC/osteonectin mRNA

50

<220>

<221> CDS

<222> (58) .. (969)

55

<223>

5 <400> 9
 cgaggagagcg cgctctgcct gccgcctgcc tgccctgccac tgagggttcc cagcacc 57

 atg agg gcc tgg atc ttc ttt ctc ctt tgc ctg gcc ggg agg gcc ttg 105
 Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu
 1 5 10 15
 10 gca gcc cct cag caa gaa gcc ctg cct gat gag aca gag gtg gtg gaa 153
 Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu
 20 25 30
 15 gaa act gtg gca gag gtg act gag gta tct gtg gga gct aat cct gtc 201
 Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val
 35 40 45
 20 cag gtg gaa gta gga gaa ttt gat gat ggt gca gag gaa acc gaa gag 249
 Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu
 50 55 60
 25 gag gtg gtg gcg gaa aat ccc tgc cag aac cac cac tgc aaa cac ggc 297
 Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly
 65 70 75 80
 30 aag gtg tgc gag ctg gat gag aac aac acc ccc atg tgc gtg tgc cag 345
 Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln
 85 90 95
 35 gac ccc acc agc tgc cca gcc ccc att ggc gag ttt gag aag gtg tgc 393
 Asp Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys
 100 105 110
 40 agc aat gac aac aag acc ttc gac tct tcc tgc cac ttc ttt gcc aca 441
 Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr
 115 120 125
 45 aag tgc acc ctg gag ggc acc aag aag ggc cac aag ctc cac ctg gac 489
 Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp
 130 135 140
 50 tac atc ggg cct tgc aaa tac atc ccc cct tgc ctg gac tct gag ctg 537
 Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu
 145 150 155 160
 55 acc gaa ttc ccc ctg cgc atg cgg gac tgg ctc aag aac gtc ctg gtc 585
 Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val
 165 170 175
 60 acc ctg tat gag agg gat gag gac aac aac ctt ctg act gag aag cag 633
 Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln
 180 185 190
 65 aag ctg cgg gtg aag aag atc cat gag aat gag aag cgc ctg gag gca 681
 Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala

195 200 205

5 gga gac cac ccc gtg gag ctg ctg gcc cgg gac ttc gag aag aac tat 729
Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr
210 215 220

10 aac atg tac atc ttc cct gta cac tgg cag ttc ggc cag ctg gac cag 777
Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln
225 230 235 240

15 cac ccc att gac ggg tac ctc tcc cac acc gag ctg gct cca ctg cgt 825
His Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg
245 250 255

20 gct ccc ctc atc ccc atg gag cat tgc acc acc cgc ttt ttc gag acc 873
Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr
260 265 270

25 tgt gac ctg gac aat gac aag tac atc gcc ctg gat gag tgg gcc ggc 921
Cys Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Asp Glu Trp Ala Gly
275 280 285

30 tgc ttc ggc atc aag cag aag gat atc gac aag gat ctt gtg atc taa 969
Cys Phe Gly Ile Lys Gln Lys Asp Ile Asp Lys Asp Leu Val Ile
290 295 300

35 atccactcct tccacagtac cggattctct cttaaacct ccccttcgtg tttccccaa 1029

40 tgtttaaaat gtttgatgg tttgttggtc tgcttgaga caaggtgcta acatagattt 1089

aagtgaatac attaacggtg ctaaaaatga aaattctaac ccaagacatg acattcttag 1149

ctgtaactta actattaagg cttttccac acgcattaat agtcccattt ttctcttgcc 1209

45 atttgtagct ttgcccattg tcttattggc acatgggtgg acacggatct gctgggctct 1269

gccttaaaca cacattgcag cttcaacttt tctctttagt gttctgtttg aaactaatac 1329

50 ttaccgagtc agactttgtg ttcatttcat ttcagggtct tggctgctg tgggcttccc 1389

cagggtggcct ggaggtgggc aaaggaagt aacagacaca cgatgtgtc aaggatgggt 1449

ttgggactag aggtcagtg gtgggagaga tccctgcaga atccaccaac cagaacgtgg 1509

55 tttgcctgag gctgtaactg agagaaagat tctggggctg tcttatgaaa atatagacat 1569

tctcacataa gccagttca tcaccatttc ctcttttacc tttcagtgc gtttcttttc 1629

acattaggct gttggttcaa acttttgga gcacggactg tcagttctct gggaagtgg 1689

cagcgcaccc tgcagggtt ctctctctct gtcttttga gaaccagggc tcttctcagg 1749

ggctctaggg actgccaggc tgtttcagcc aggaaggcca aaatcaagag tgagatgtag 1809

aaagttgtaa aatagaaaaa gtggagttgg tgaatcgggt gttctttcct cacatttgga 1869

tgattgtcat aaggttttta gcatgttcct ctttttcttc accctcccct ttgtttcttct 1929
attaatcaag agaaacttca aagttaatgg gatggtcgga tctcacaggc tgagaactcg 1989
5 ttcacctcca agcatttcat gaaaaagctg cttcttatta atcatacaaa ctctcaccat 2049
gatgtgaaga gtttcacaaa tctttcaaaa taaaaagtaa tgacttagaa actgaaaaaa 2109
aaaaaaaaa aaaaaaaaaa aaaa 2133

<210> 10

<211> 303

<212> PRT

<213> Homo sapiens

<220>

<221> SIGNAL

<222> (1)..(17)

<223>

<220>

<221> gene

<222> (1)..(303)

<223> SPARC, osteonectin

<220>

<221> CHAIN

<222> (18)..(303)

<223> SPARC

<220>

<221> DOMAIN

<222> (22)..(69)

<223> ASP/GLU-RICH (ACIDIC, BINDS CALCIUM)

5

<220>

<221> DOMAIN

10

<222> (94)..(149)

<223> KAZAL-LIKE

15

<220>

<221> DOMAIN

20

<222> (185)..(221)

<223> ALPHA-HELIX CONFORMATION (PROBABLE)

25

<220>

<221> CA_BIND

30

<222> (274)..(285)

<223> 12 EF-HAND

35

<220>

<221> CARBOHYD

40

<222> (116)..(116)

<223> N-LINKED (GLCNAC ...) (PROBABLE)

45

<220>

<221> DISULFID

50

<222> (72)..(83)

<223>

55

<220>

<221> DISULFID

<222> (77) .. (93)

<223>

<220>

<221> DISULFID

<222> (95) .. (130)

<223>

<220>

<221> DISULFID

<222> (101) .. (123)

<223>

<220>

<221> DISULFID

<222> (112) .. (149)

<223>

<220>

<221> DISULFID

<222> (155) .. (265)

<223>

<220>

<221> DISULFID

<222> (273) .. (289)

<223>

<400> 10

5 Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu
1 5 10 15

10 Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu
20 25 30

15 Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val
35 40 45

20 Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu
50 55 60

25 Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly
65 70 75 80

30 Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln
85 90 95

35 Asp Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys
100 105 110

40 Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr
115 120 125

45 Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp
130 135 140

50 Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu
145 150 155 160

55 Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val
165 170 175

Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln
180 185 190

Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala
195 200 205

Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr
210 215 220

5

Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln
225 230 235 240

10

His Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg
245 250 255

15

Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr
260 265 270

20

Cys Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Asp Glu Trp Ala Gly
275 280 285

25

Cys Phe Gly Ile Lys Gln Lys Asp Ile Asp Lys Asp Leu Val Ile
290 295 300

30

<210> 11
<211> 4586
<212> DNA
<213> Homo sapiens

35

<220>
<221> gene

40

<222> (1) .. (4586)
<223> mRNA for met proto-oncogene

45

<220>
<221> REFERENCE

50

<222> (1) .. (4586)
<223> Park, M., Dean, M., Kaul, K., Braun, M.J., Gonda, M.A. and Vande Woude, G. TITLE Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (18), 6379-6383 (1987)

55

<220>

5 <221> CDS

<222> (195)..(4367)

<223>

10

<400> 11

15

gaattccgcc ctgcgcgcc gccggcgcccc gagcgctttg tgagcagatg cggagccgag 60

tggagggcgc gagccagatg cggggcgaca gctgacttgc tgagaggagg cggggaggcg 120

cggagcgcgc gtgtggtcct tgcgcgcgtg acttctccac tggttcctgg gcaccgaaag 180

20

ataaacctct cata atg aag gcc ccc gct gtg ctt gca cct ggc atc ctc 230
Met Lys Ala Pro Ala Val Leu Ala Pro Gly Ile Leu
1 5 10

25

gtg ctc ctg ttt acc ttg gtg cag agg agc aat ggg gag tgt aaa gag 278
Val Leu Leu Phe Thr Leu Val Gln Arg Ser Asn Gly Glu Cys Lys Glu
15 20 25

30

gca cta gca aag tcc gag atg aat gtg aat atg aag tat cag ctt ccc 326
Ala Leu Ala Lys Ser Glu Met Asn Val Asn Met Lys Tyr Gln Leu Pro
30 35 40

35

aac ttc acc gcg gaa aca ccc atc cag aat gtc att cta cat gag cat 374
Asn Phe Thr Ala Glu Thr Pro Ile Gln Asn Val Ile Leu His Glu His
45 50 55 60

cac att ttc ctt ggt gcc act aac tac att tat gtt tta aat gag gaa 422
His Ile Phe Leu Gly Ala Thr Asn Tyr Ile Tyr Val Leu Asn Glu Glu
65 70 75

40

gac ctt cag aag gtt gct gag tac aag act ggg cct gtg ctg gaa cac 470
Asp Leu Gln Lys Val Ala Glu Tyr Lys Thr Gly Pro Val Leu Glu His
80 85 90

45

cca gat tgt ttc cca tgt cag gac tgc agc agc aaa gcc aat tta tca 518
Pro Asp Cys Phe Pro Cys Gln Asp Cys Ser Ser Lys Ala Asn Leu Ser
95 100 105

50

gga ggt gtt tgg aaa gat aac atc aac atg gct cta gtt gtc gac acc 566
Gly Gly Val Trp Lys Asp Asn Ile Asn Met Ala Leu Val Val Asp Thr
110 115 120

55

tac tat gat gat caa ctc att agc tgt ggc agc gtc aac aga ggg acc 614
Tyr Tyr Asp Asp Gln Leu Ile Ser Cys Gly Ser Val Asn Arg Gly Thr
125 130 135 140

tgc cag cga cat gtc ttt ccc cac aat cat act gct gac ata cag tcg 662

	Cys	Gln	Arg	His	Val	Phe	Pro	His	Asn	His	Thr	Ala	Asp	Ile	Gln	Ser	
					145					150					155		
5	gag	gtt	cac	tgc	ata	ttc	tcc	cca	cag	ata	gaa	gag	ccc	agc	cag	tgt	710
	Glu	Val	His	Cys	Ile	Phe	Ser	Pro	Gln	Ile	Glu	Glu	Pro	Ser	Gln	Cys	
				160					165					170			
10	cct	gac	tgt	gtg	gtg	agc	gcc	ctg	gga	gcc	aaa	gtc	ctt	tca	tct	gta	758
	Pro	Asp	Cys	Val	Val	Ser	Ala	Leu	Gly	Ala	Lys	Val	Leu	Ser	Ser	Val	
				175				180						185			
15	aag	gac	cgg	ttc	atc	aac	ttc	ttt	gta	ggc	aat	acc	ata	aat	tct	tct	806
	Lys	Asp	Arg	Phe	Ile	Asn	Phe	Phe	Val	Gly	Asn	Thr	Ile	Asn	Ser	Ser	
				190			195					200					
20	tat	ttc	cca	gat	cat	cca	ttg	cat	tcg	ata	tca	gtg	aga	agg	cta	aag	854
	Tyr	Phe	Pro	Asp	His	Pro	Leu	His	Ser	Ile	Ser	Val	Arg	Arg	Leu	Lys	
				205			210				215					220	
25	gaa	acg	aaa	gat	ggg	ttt	atg	ttt	ttg	acg	gac	cag	tcc	tac	att	gat	902
	Glu	Thr	Lys	Asp	Gly	Phe	Met	Phe	Leu	Thr	Asp	Gln	Ser	Tyr	Ile	Asp	
					225					230					235		
30	gtt	tta	cct	gag	ttc	aga	gat	tct	tac	ccc	att	aag	tat	gtc	cat	gcc	950
	Val	Leu	Pro	Glu	Phe	Arg	Asp	Ser	Tyr	Pro	Ile	Lys	Tyr	Val	His	Ala	
				240					245					250			
35	ttt	gaa	agc	aac	aat	ttt	att	tac	ttc	ttg	acg	gtc	caa	agg	gaa	act	998
	Phe	Glu	Ser	Asn	Asn	Phe	Ile	Tyr	Phe	Leu	Thr	Val	Gln	Arg	Glu	Thr	
				255				260					265				
40	cta	gat	gct	cag	act	ttt	cac	aca	aga	ata	atc	agg	ttc	tgt	tcc	ata	1046
	Leu	Asp	Ala	Gln	Thr	Phe	His	Thr	Arg	Ile	Ile	Arg	Phe	Cys	Ser	Ile	
				270			275					280					
45	aac	tct	gga	ttg	cat	tcc	tac	atg	gaa	atg	cct	ctg	gag	tgt	att	ctc	1094
	Asn	Ser	Gly	Leu	His	Ser	Tyr	Met	Glu	Met	Pro	Leu	Glu	Cys	Ile	Leu	
				285			290				295					300	
50	aca	gaa	aag	aga	aaa	aag	aga	tcc	aca	aag	aag	gaa	gtg	ttt	aat	ata	1142
	Thr	Glu	Lys	Arg	Lys	Lys	Arg	Ser	Thr	Lys	Lys	Glu	Val	Phe	Asn	Ile	
					305					310					315		
55	ctt	cag	gct	gcg	tat	gtc	agc	aag	cct	ggg	gcc	cag	ctt	gct	aga	caa	1190
	Leu	Gln	Ala	Ala	Tyr	Val	Ser	Lys	Pro	Gly	Ala	Gln	Leu	Ala	Arg	Gln	
				320					325					330			
60	ata	gga	gcc	agc	ctg	aat	gat	gac	att	ctt	ttc	ggg	gtg	ttc	gca	caa	1238
	Ile	Gly	Ala	Ser	Leu	Asn	Asp	Asp	Ile	Leu	Phe	Gly	Val	Phe	Ala	Gln	
				335			340						345				
65	agc	aag	cca	gat	tct	gcc	gaa	cca	atg	gat	cga	tct	gcc	atg	tgt	gca	1286
	Ser	Lys	Pro	Asp	Ser	Ala	Glu	Pro	Met	Asp	Arg	Ser	Ala	Met	Cys	Ala	
				350			355				360						
70	ttc	cct	atc	aaa	tat	gtc	aac	gac	ttc	ttc	aac	aag	atc	gtc	aac	aaa	1334

5
10
15
20
25
30
35
40
45
50
55

Phe Pro Ile Lys Tyr Val Asn Asp Phe Phe Asn Lys Ile Val Asn Lys
365 370 375 380

aac aat gtg aga tgt ctc cag cat ttt tac gga ccc aat cat gag cac 1382
Asn Asn Val Arg Cys Leu Gln His Phe Tyr Gly Pro Asn His Glu His
385 390 395

tgc ttt aat agg aca ctt ctg aga aat tca tca ggc tgt gaa gcg cgc 1430
Cys Phe Asn Arg Thr Leu Leu Arg Asn Ser Ser Gly Cys Glu Ala Arg
400 405 410

cgt gat gaa tat cga aca gag ttt acc aca gct ttg cag cgc gtt gac 1478
Arg Asp Glu Tyr Arg Thr Glu Phe Thr Thr Ala Leu Gln Arg Val Asp
415 420 425

tta ttc atg ggt caa ttc agc gaa gtc ctc tta aca tct ata tcc acc 1526
Leu Phe Met Gly Gln Phe Ser Glu Val Leu Leu Thr Ser Ile Ser Thr
430 435 440

ttc att aaa gga gac ctc acc ata gct aat ctt ggg aca tca gag ggt 1574
Phe Ile Lys Gly Asp Leu Thr Ile Ala Asn Leu Gly Thr Ser Glu Gly
445 450 455 460

cgc ttc atg cag gtt gtg gtt tct cga tca gga cca tca acc cct cat 1622
Arg Phe Met Gln Val Val Ser Arg Ser Gly Pro Ser Thr Pro His
465 470 475

gtg aat ttt ctc ctg gac tcc cat cca gtg tct cca gaa gtg att gtg 1670
Val Asn Phe Leu Leu Asp Ser His Pro Val Ser Pro Glu Val Ile Val
480 485 490

gag cat aca tta aac caa aat ggc tac aca ctg gtt atc act ggg aag 1718
Glu His Thr Leu Asn Gln Asn Gly Tyr Thr Leu Val Ile Thr Gly Lys
495 500 505

aag atc acg aag atc cca ttg aat ggc ttg ggc tgc aga cat ttc cag 1766
Lys Ile Thr Lys Ile Pro Leu Asn Gly Leu Gly Cys Arg His Phe Gln
510 515 520

tcc tgc agt caa tgc ctc tct gcc cca ccc ttt gtt cag tgt ggc tgg 1814
Ser Cys Ser Gln Cys Leu Ser Ala Pro Pro Phe Val Gln Cys Gly Trp
525 530 535 540

tgc cac gac aaa tgt gtg cga tgc gag gaa tgc ctg agc ggg aca tgg 1862
Cys His Asp Lys Cys Val Arg Ser Glu Glu Cys Leu Ser Gly Thr Trp
545 550 555

act caa cag atc tgt ctg cct gca atc tac aag gtt ttc cca aat agt 1910
Thr Gln Gln Ile Cys Leu Pro Ala Ile Tyr Lys Val Phe Pro Asn Ser
560 565 570

gca ccc ctt gaa gga ggg aca agg ctg acc ata tgt ggc tgg gac ttt 1958
Ala Pro Leu Glu Gly Gly Thr Arg Leu Thr Ile Cys Gly Trp Asp Phe
575 580 585

gga ttt cgg agg aat aat aaa ttt gat tta aag aaa act aga gtt ctc 2006

5
10
15
20
25
30
35
40
45
50
55

Gly	Phe	Arg	Arg	Asn	Asn	Lys	Phe	Asp	Leu	Lys	Lys	Thr	Arg	Val	Leu	
590						595					600					
ctt	gga	aat	gag	agc	tgc	acc	ttg	act	tta	agt	gag	agc	acg	atg	aat	2054
Leu	Gly	Asn	Glu	Ser	Cys	Thr	Leu	Thr	Leu	Ser	Glu	Ser	Thr	Met	Asn	
605					610					615					620	
aca	ttg	aaa	tgc	aca	gtt	ggc	cct	gcc	atg	aat	aag	cat	ttc	aat	atg	2102
Thr	Leu	Lys	Cys	Thr	Val	Gly	Pro	Ala	Met	Asn	Lys	His	Phe	Asn	Met	
				625					630						635	
tcc	ata	att	att	tca	aat	ggc	cac	ggg	aca	aca	caa	tac	agt	aca	ttc	2150
Ser	Ile	Ile	Ile	Ser	Asn	Gly	His	Gly	Thr	Thr	Gln	Tyr	Ser	Thr	Phe	
			640					645						650		
tcc	tat	gtg	gat	cct	gta	ata	aca	agt	att	tcg	ccg	aaa	tac	ggc	cct	2198
Ser	Tyr	Val	Asp	Pro	Val	Ile	Thr	Ser	Ile	Ser	Pro	Lys	Tyr	Gly	Pro	
		655					660					665				
atg	gct	ggc	ggc	act	tta	ctt	act	tta	act	gga	aat	tac	cta	aac	agt	2246
Met	Ala	Gly	Gly	Thr	Leu	Leu	Thr	Leu	Thr	Gly	Asn	Tyr	Leu	Asn	Ser	
	670					675					680					
ggg	aat	tct	aga	cac	att	tca	att	ggc	gga	aaa	aca	tgt	act	tta	aaa	2294
Gly	Asn	Ser	Arg	His	Ile	Ser	Ile	Gly	Gly	Lys	Thr	Cys	Thr	Leu	Lys	
685					690					695					700	
agt	gtg	tca	aac	agt	att	ctt	gaa	tgt	tat	acc	cca	gcc	caa	acc	att	2342
Ser	Val	Ser	Asn	Ser	Ile	Leu	Glu	Cys	Tyr	Thr	Pro	Ala	Gln	Thr	Ile	
				705					710						715	
tca	act	gag	ttt	gct	gtt	aaa	ttg	aaa	att	gac	tta	gcc	aac	cga	gag	2390
Ser	Thr	Glu	Phe	Ala	Val	Lys	Leu	Lys	Ile	Asp	Leu	Ala	Asn	Arg	Glu	
			720					725						730		
aca	agc	atc	ttc	agt	tac	cgt	gaa	gat	ccc	att	gtc	tat	gaa	att	cat	2438
Thr	Ser	Ile	Phe	Ser	Tyr	Arg	Glu	Asp	Pro	Ile	Val	Tyr	Glu	Ile	His	
		735					740					745				
cca	acc	aaa	tct	ttt	att	agt	ggc	ggg	agc	aca	ata	aca	ggc	gtt	ggg	2486
Pro	Thr	Lys	Ser	Phe	Ile	Ser	Gly	Gly	Ser	Thr	Ile	Thr	Gly	Val	Gly	
	750					755					760					
aaa	aac	ctg	aat	tca	gtt	agt	gtc	ccg	aga	atg	gtc	ata	aat	gtg	cat	2534
Lys	Asn	Leu	Asn	Ser	Val	Ser	Val	Pro	Arg	Met	Val	Ile	Asn	Val	His	
765					770					775					780	
gaa	gca	gga	agg	aac	ttt	aca	gtg	gca	tgt	caa	cat	cgc	tct	aat	tca	2582
Glu	Ala	Gly	Arg	Asn	Phe	Thr	Val	Ala	Cys	Gln	His	Arg	Ser	Asn	Ser	
				785					790						795	
gag	ata	atc	tgt	tgt	acc	act	cct	tcc	ctg	caa	cag	ctg	aat	ctg	caa	2630
Glu	Ile	Ile	Cys	Cys	Thr	Thr	Pro	Ser	Leu	Gln	Gln	Leu	Asn	Leu	Gln	
			800					805						810		
ctc	ccc	ctg	aaa	acc	aaa	gcc	ttt	ttc	atg	tta	gat	ggg	atc	ctt	tcc	2678

Leu Pro Leu Lys Thr Lys Ala Phe Phe Met Leu Asp Gly Ile Leu Ser
815 820 825

5 aaa tac ttt gat ctc att tat gta cat aat cct gtg ttt aag cct ttt 2726
Lys Tyr Phe Asp Leu Ile Tyr Val His Asn Pro Val Phe Lys Pro Phe
830 835 840

10 gaa aag cca gtg atg atc tca atg ggc aat gaa aat gta ctg gaa att 2774
Glu Lys Pro Val Met Ile Ser Met Gly Asn Glu Asn Val Leu Glu Ile
845 850 855 860

15 aag gga aat gat att gac cct gaa gca gtt aaa ggt gaa gtg tta aaa 2822
Lys Gly Asn Asp Ile Asp Pro Glu Ala Val Lys Gly Glu Val Leu Lys
865 870 875

20 gtt gga aat aag agc tgt gag aat ata cac tta cat tct gaa gcc gtt 2870
Val Gly Asn Lys Ser Cys Glu Asn Ile His Leu His Ser Glu Ala Val
880 885 890

25 tta tgc acg gtc ccc aat gac ctg ctg aaa ttg aac agc gag cta aat 2918
Leu Cys Thr Val Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu Asn
895 900 905

30 ata gag tgg aag caa gca att tct tca acc gtc ctt gga aaa gta ata 2966
Ile Glu Trp Lys Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val Ile
910 915 920

35 gtt caa cca gat cag aat ttc aca gga ttg att gct ggt gtt gtc tca 3014
Val Gln Pro Asp Gln Asn Phe Thr Gly Leu Ile Ala Gly Val Val Ser
925 930 935 940

40 ata tca aca gca ctg tta tta cta ctt ggg ttt ttc ctg tgg ctg aaa 3062
Ile Ser Thr Ala Leu Leu Leu Leu Leu Gly Phe Phe Leu Trp Leu Lys
945 950 955

45 aag aga aag caa att aaa gat ctg ggc agt gaa tta gtt cgc tac gat 3110
Lys Arg Lys Gln Ile Lys Asp Leu Gly Ser Glu Leu Val Arg Tyr Asp
960 965 970

50 gca aga gta cac act cct cat ttg gat agg ctt gta agt gcc cga agt 3158
Ala Arg Val His Thr Pro His Leu Asp Arg Leu Val Ser Ala Arg Ser
975 980 985

55 gta agc cca act aca gaa atg gtt tca aat gaa tct gta gac tac cga 3206
Val Ser Pro Thr Thr Glu Met Val Ser Asn Glu Ser Val Asp Tyr Arg
990 995 1000

60 gct act ttt cca gaa gat cag ttt cct aat tca tct cag aac ggt 3251
Ala Thr Phe Pro Glu Asp Gln Phe Pro Asn Ser Ser Gln Asn Gly
1005 1010 1015

65 tca tgc cga caa gtg cag tat cct ctg aca gac atg tcc ccc atc 3296
Ser Cys Arg Gln Val Gln Tyr Pro Leu Thr Asp Met Ser Pro Ile
1020 1025 1030

70 cta act agt ggg gac tct gat ata tcc agt cca tta ctg caa aat 3341

5 Leu Thr Ser Gly Asp Ser Asp Ile Ser Ser Pro Leu Leu Gln Asn
 1035 1040 1045
 act gtc cac att gac ctc agt gct cta aat cca gag ctg gtc cag 3386
 Thr Val His Ile Asp Leu Ser Ala Leu Asn Pro Glu Leu Val Gln
 1050 1055 1060
 10 gca gtg cag cat gta gtg att ggg ccc agt agc ctg att gtg cat 3431
 Ala Val Gln His Val Val Ile Gly Pro Ser Ser Leu Ile Val His
 1065 1070 1075
 15 ttc aat gaa gtc ata gga aga ggg cat ttt ggt tgt gta tat cat 3476
 Phe Asn Glu Val Ile Gly Arg Gly His Phe Gly Cys Val Tyr His
 1080 1085 1090
 20 ggg act ttg ttg gac aat gat ggc aag aaa att cac tgt gct gtg 3521
 Gly Thr Leu Leu Asp Asn Asp Gly Lys Lys Ile His Cys Ala Val
 1095 1100 1105
 25 aaa tcc ttg aac aga atc act gac ata gga gaa gtt tcc caa ttt 3566
 Lys Ser Leu Asn Arg Ile Thr Asp Ile Gly Glu Val Ser Gln Phe
 1110 1115 1120
 30 ctg acc gag gga atc atc atg aaa gat ttt agt cat ccc aat gtc 3611
 Leu Thr Glu Gly Ile Ile Met Lys Asp Phe Ser His Pro Asn Val
 1125 1130 1135
 35 ctc tcg ctc ctg gga atc tgc ctg cga agt gaa ggg tct ccg ctg 3656
 Leu Ser Leu Leu Gly Ile Cys Leu Arg Ser Glu Gly Ser Pro Leu
 1140 1145 1150
 40 gtg gtc cta cca tac atg aaa cat gga gat ctt cga aat ttc att 3701
 Val Val Leu Pro Tyr Met Lys His Gly Asp Leu Arg Asn Phe Ile
 1155 1160 1165
 45 cga aat gag act cat aat cca act gta aaa gat ctt att ggc ttt 3746
 Arg Asn Glu Thr His Asn Pro Thr Val Lys Asp Leu Ile Gly Phe
 1170 1175 1180
 50 ggt ctt caa gta gcc aaa ggc atg aaa tat ctt gca agc aaa aag 3791
 Gly Leu Gln Val Ala Lys Gly Met Lys Tyr Leu Ala Ser Lys Lys
 1185 1190 1195
 55 ttt gtc cac aga gac ttg gct gca aga aac tgt atg ctg gat gaa 3836
 Phe Val His Arg Asp Leu Ala Ala Arg Asn Cys Met Leu Asp Glu
 1200 1205 1210
 60 aaa ttc aca gtc aag gtt gct gat ttt ggt ctt gcc aga gac atg 3881
 Lys Phe Thr Val Lys Val Ala Asp Phe Gly Leu Ala Arg Asp Met
 1215 1220 1225
 65 tat gat aaa gaa tac tat agt gta cac aac aaa aca ggt gca aag 3926
 Tyr Asp Lys Glu Tyr Tyr Ser Val His Asn Lys Thr Gly Ala Lys
 1230 1235 1240
 70 ctg cca gtg aag tgg atg gct ttg gaa agt ctg caa act caa aag 3971

Leu Pro Val Lys Trp Met Ala Leu Glu Ser Leu Gln Thr Gln Lys
1245 1250 1255

5 ttt acc acc aag tca gat gtg tgg tcc ttt ggc gtc gtc ctc tgg 4016
Phe Thr Thr Lys Ser Asp Val Trp Ser Phe Gly Val Val Leu Trp
1260 1265 1270

10 gag ctg atg aca aga gga gcc cca cct tat cct gac gta aac acc 4061
Glu Leu Met Thr Arg Gly Ala Pro Pro Tyr Pro Asp Val Asn Thr
1275 1280 1285

15 ttt gat ata act gtt tac ttg ttg caa ggg aga aga ctc cta caa 4106
Phe Asp Ile Thr Val Tyr Leu Leu Gln Gly Arg Arg Leu Leu Gln
1290 1295 1300

20 ccc gaa tac tgc cca gac ccc tta tat gaa gta atg cta aaa tgc 4151
Pro Glu Tyr Cys Pro Asp Pro Leu Tyr Glu Val Met Leu Lys Cys
1305 1310 1315

25 tgg cac cct aaa gcc gaa atg cgc cca tcc ttt tct gaa ctg gtg 4196
Trp His Pro Lys Ala Glu Met Arg Pro Ser Phe Ser Glu Leu Val
1320 1325 1330

30 tcc cgg ata tca gcg atc ttc tct act ttc att ggg gag cac tat 4241
Ser Arg Ile Ser Ala Ile Phe Ser Thr Phe Ile Gly Glu His Tyr
1335 1340 1345

35 gtc cat gtg aac gct act tat gtg aac gta aaa tgt gtc gct ccg 4286
Val His Val Asn Ala Thr Tyr Val Asn Val Lys Cys Val Ala Pro
1350 1355 1360

40 tat cct tct ctg ttg tca tca gaa gat aac gct gat gat gag gtg 4331
Tyr Pro Ser Leu Leu Ser Ser Glu Asp Asn Ala Asp Asp Glu Val
1365 1370 1375

45 gac aca cga cca gcc tcc ttc tgg gag aca tca tag tgctagtact 4377
Asp Thr Arg Pro Ala Ser Phe Trp Glu Thr Ser
1380 1385 1390

atgtcaaagc aacagtccac actttgtcca atggtttttt cactgcctga cctttaaaag 4437

gccatcgata ttcttttgctc cttgccaaat tgcactatta ataggacttg tattggttatt 4497

45 taaattactg gattctaagg aatttcttat ctgacagagc atcagaacca gaggcttggt 4557

cccacaggcc agggaccaat gcgctgcag 4586

50 <210> 12

<211> 1390

<212> PRT

55 <213> Homo sapiens

<220>

5 <221> SIGNAL

<222> (1) .. (24)

<223>

10

<220>

15 <221> gene

<222> (1) .. (1390)

<223> C-Met Proto Oncogene

20

<220>

25 <221> CHAIN

<222> (25) .. (1390)

<223> HEPATOCYTE GROWTH FACTOR RECEPTOR

30

<220>

35 <221> DOMAIN

<222> (25) .. (932)

<223> EXTRACELLULAR (POTENTIAL)

40

<220>

45 <221> TRANSMEM

<222> (933) .. (955)

<223> POTENTIAL

50

<220>

55 <221> DOMAIN

<222> (956)..(1390)

<223> CYTOPLASMIC (POTENTIAL)

5

<220>

<221> DOMAIN

10

<222> (55)..(500)

<223> SEMA.

15

<220>

<221> DOMAIN

20

<222> (1078)..(1345)

<223> PROTEIN KINASE

25

<220>

<221> SITE

30

<222> (1009)..(1010)

<223> Breakpoint for Translocation to Form TPR-Met Oncogene

35

<220>

<221> NP_BIND

40

<222> (1084)..(1092)

<223> ATP (BY SIMILARITY)

45

<220>

<221> BINDING

50

<222> (1110)..(1110)

<223> ATP (BY SIMILARITY)

55

FOOTNOTES

<220>

<221> ACT_SITE

5 <222> (1204)..(1204)

<223> BY SIMILARITY

10

<220>

<221> SITE

15 <222> (307)..(308)

<223> CLEAVAGE (POTENTIAL)

20

<220>

<221> MOD_RES

25 <222> (1235)..(1235)

<223> PHOSPHORYLATION (AUTO-)

30

<220>

<221> CARBOHYD

35 <222> (45)..(45)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

40

<220>

<221> CARBOHYD

45 <222> (106)..(106)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

50

<220>

<221> CARBOHYD

55 <222> (149)..(149)

TOPT-00000000

<223> N-LINKED (GLCNAC...) (POTENTIAL)

5 <220>

<221> CARBOHYD

<222> (202)..(202)

10 <223> N-LINKED (GLCNAC...) (POTENTIAL)

15 <220>

<221> CARBOHYD

<222> (399)..(399)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

25 <220>

<221> CARBOHYD

<222> (405)..(405)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

35 <220>

<221> CARBOHYD

<222> (607)..(607)

40 <223> N-LINKED (GLCNAC...) (POTENTIAL)

45 <220>

<221> CARBOHYD

<222> (635)..(635)

50 <223> N-LINKED (GLCNAC...) (POTENTIAL)

55 <220>

<221> CARBOHYD

<222> (785)..(785)

5 <223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

10

<221> CARBOHYD

<222> (879)..(879)

15 <223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

20

<221> CARBOHYD

<222> (930)..(930)

25 <223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

30

<221> VARSPLIC

<222> (755)..(755)

35 <223> S -> STWWKEPLNIVSFLFCFAS (IN ISOFORM 2)

<220>

40

<221> VARIANT

<222> (320)..(320)

45 <223> A -> V

<220>

50

<221> VARIANT

<222> (1131)..(1131)

55 <223> M -> T (IN HPRC; GERMLINE MUTATION)

<220>

5 <221> VARIANT

<222> (1188)..(1188)

10 <223> V -> L (IN HPRC; GERMLINE MUTATION)

<220>

15 <221> VARIANT

<222> (1195)..(1195)

20 <223> L -> V (IN HPRC; GERMLINE MUTATION)

<220>

25 <221> VARIANT

<222> (1220)..(1220)

30 <223> V -> I (IN HPRC; GERMLINE MUTATION)

<220>

35 <221> VARIANT

<222> (1228)..(1228)

40 <223> D -> N (IN HPRC; GERMLINE MUTATION)

<220>

45 <221> VARIANT

<222> (1228)..(1228)

50 <223> D -> H (IN HPRC; GERMLINE MUTATION)

<220>

55 <221> VARIANT

<222> (1230)..(1230)

<223> Y -> C (IN HPRC; GERMLINE MUTATION)

5

<220>

<221> VARIANT

10

<222> (1230)..(1230)

<223> Y -> H (IN HPRC; GERMLINE MUTATION)

15

<220>

<221> VARIANT

20

<222> (1250)..(1250)

<223> M -> T (IN HPRC; GERMLINE MUTATION)

25

<400> 12

Met Lys Ala Pro Ala Val Leu Ala Pro Gly Ile Leu Val Leu Leu Phe
1 5 10 15

Thr Leu Val Gln Arg Ser Asn Gly Glu Cys Lys Glu Ala Leu Ala Lys
20 25 30

Ser Glu Met Asn Val Asn Met Lys Tyr Gln Leu Pro Asn Phe Thr Ala
35 40 45

Glu Thr Pro Ile Gln Asn Val Ile Leu His Glu His His Ile Phe Leu
50 55 60

Gly Ala Thr Asn Tyr Ile Tyr Val Leu Asn Glu Glu Asp Leu Gln Lys
65 70 75 80

Val Ala Glu Tyr Lys Thr Gly Pro Val Leu Glu His Pro Asp Cys Phe
85 90 95

Pro Cys Gln Asp Cys Ser Ser Lys Ala Asn Leu Ser Gly Gly Val Trp
100 105 110

55

5 Gln Leu Ile Ser Cys Gly Ser Val Asn Arg Gly Thr Cys Gln Arg His
 130 135 140

Ile Phe Ser Pro Gln Ile Glu Glu Pro Ser Gln Cys Pro Asp Cys Val
165 170 175

20
Ile Asn Phe Phe Val Gly Asn Thr Ile Asn Ser Ser Tyr Phe Pro Asp
195 200 205

30 Gly Phe Met Phe Leu Thr Asp Gln Ser Tyr Ile Asp Val Leu Pro Glu
225 230 235 240

Asn Phe Ile Tyr Phe Leu Thr Val Gln Arg Glu Thr Leu Asp Ala Gln
260 265 270

45 His Ser Tyr Met Glu Met Pro Leu Glu Cys Ile Leu Thr Glu Lys Arg
290 295 300

Tyr Val Ser Lys Pro Gly Ala Gln Leu Ala Arg Gln Ile Gly Ala Ser
325 330 335

Leu Asn Asp Asp Ile Leu Phe Gly Val Phe Ala Gln Ser Lys Pro Asp
340 345 350

5 Ser Ala Glu Pro Met Asp Arg Ser Ala Met Cys Ala Phe Pro Ile Lys
355 360 365

10 Tyr Val Asn Asp Phe Phe Asn Lys Ile Val Asn Lys Asn Asn Val Arg
370 375 380

15 Cys Leu Gln His Phe Tyr Gly Pro Asn His Glu His Cys Phe Asn Arg
385 390 395 400

20 Thr Leu Leu Arg Asn Ser Ser Gly Cys Glu Ala Arg Arg Asp Glu Tyr
405 410 415

25 Arg Thr Glu Phe Thr Thr Ala Leu Gln Arg Val Asp Leu Phe Met Gly
420 425 430

30 Gln Phe Ser Glu Val Leu Leu Thr Ser Ile Ser Thr Phe Ile Lys Gly
435 440 445

35 Asp Leu Thr Ile Ala Asn Leu Gly Thr Ser Glu Gly Arg Phe Met Gln
450 455 460

40 Val Val Val Ser Arg Ser Gly Pro Ser Thr Pro His Val Asn Phe Leu
465 470 475 480

45 Leu Asp Ser His Pro Val Ser Pro Glu Val Ile Val Glu His Thr Leu
485 490 495

50 Asn Gln Asn Gly Tyr Thr Leu Val Ile Thr Gly Lys Lys Ile Thr Lys
500 505 510

55 Ile Pro Leu Asn Gly Leu Gly Cys Arg His Phe Gln Ser Cys Ser Gln
515 520 525

Cys Leu Ser Ala Pro Pro Phe Val Gln Cys Gly Trp Cys His Asp Lys
530 535 540

Cys Val Arg Ser Glu Glu Cys Leu Ser Gly Thr Trp Thr Gln Gln Ile
545 550 555 560

5 Cys Leu Pro Ala Ile Tyr Lys Val Phe Pro Asn Ser Ala Pro Leu Glu
 565 570 575
 Gly Gly Thr Arg Leu Thr Ile Cys Gly Trp Asp Phe Gly Phe Arg Arg
 580 585 590
 10 Asn Asn Lys Phe Asp Leu Lys Lys Thr Arg Val Leu Leu Gly Asn Glu
 595 600 605
 15 Ser Cys Thr Leu Thr Leu Ser Glu Ser Thr Met Asn Thr Leu Lys Cys
 610 615 620
 Thr Val Gly Pro Ala Met Asn Lys His Phe Asn Met Ser Ile Ile Ile
 625 630 635 640
 20 Ser Asn Gly His Gly Thr Thr Gln Tyr Ser Thr Phe Ser Tyr Val Asp
 645 650 655
 25 Pro Val Ile Thr Ser Ile Ser Pro Lys Tyr Gly Pro Met Ala Gly Gly
 660 665 670
 30 Thr Leu Leu Thr Leu Thr Gly Asn Tyr Leu Asn Ser Gly Asn Ser Arg
 675 680 685
 35 His Ile Ser Ile Gly Gly Lys Thr Cys Thr Leu Lys Ser Val Ser Asn
 690 695 700
 Ser Ile Leu Glu Cys Tyr Thr Pro Ala Gln Thr Ile Ser Thr Glu Phe
 705 710 715 720
 40 Ala Val Lys Leu Lys Ile Asp Leu Ala Asn Arg Glu Thr Ser Ile Phe
 725 730 735
 45 Ser Tyr Arg Glu Asp Pro Ile Val Tyr Glu Ile His Pro Thr Lys Ser
 740 745 750
 50 Phe Ile Ser Gly Gly Ser Thr Ile Thr Gly Val Gly Lys Asn Leu Asn
 755 760 765
 55 Ser Val Ser Val Pro Arg Met Val Ile Asn Val His Glu Ala Gly Arg
 770 775 780

Asn Phe Thr Val Ala Cys Gln His Arg Ser Asn Ser Glu Ile Ile Cys
785 790 795 800

5 Cys Thr Thr Pro Ser Leu Gln Gln Leu Asn Leu Gln Leu Pro Leu Lys
805 810 815

10 Thr Lys Ala Phe Phe Met Leu Asp Gly Ile Leu Ser Lys Tyr Phe Asp
820 825 830

15 Leu Ile Tyr Val His Asn Pro Val Phe Lys Pro Phe Glu Lys Pro Val
835 840 845

20 Met Ile Ser Met Gly Asn Glu Asn Val Leu Glu Ile Lys Gly Asn Asp
850 855 860

Ile Asp Pro Glu Ala Val Lys Gly Glu Val Leu Lys Val Gly Asn Lys
865 870 875 880

25 Ser Cys Glu Asn Ile His Leu His Ser Glu Ala Val Leu Cys Thr Val
885 890 895

30 Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu Asn Ile Glu Trp Lys
900 905 910

Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val Ile Val Gln Pro Asp
915 920 925

35 Gln Asn Phe Thr Gly Leu Ile Ala Gly Val Val Ser Ile Ser Thr Ala
930 935 940

40 Leu Leu Leu Leu Leu Gly Phe Phe Leu Trp Leu Lys Lys Arg Lys Gln
945 950 955 960

45 Ile Lys Asp Leu Gly Ser Glu Leu Val Arg Tyr Asp Ala Arg Val His
965 970 975

50 Thr Pro His Leu Asp Arg Leu Val Ser Ala Arg Ser Val Ser Pro Thr
980 985 990

55 Thr Glu Met Val Ser Asn Glu Ser Val Asp Tyr Arg Ala Thr Phe Pro
995 1000 1005

5
10
15
20
25
30
35
40
45
50
55

Glu Asp Gln Phe Pro Asn Ser Ser Gln Asn Gly Ser Cys Arg Gln
 1010 1015 1020
 Val Gln Tyr Pro Leu Thr Asp Met Ser Pro Ile Leu Thr Ser Gly
 1025 1030 1035
 Asp Ser Asp Ile Ser Ser Pro Leu Leu Gln Asn Thr Val His Ile
 1040 1045 1050
 Asp Leu Ser Ala Leu Asn Pro Glu Leu Val Gln Ala Val Gln His
 1055 1060 1065
 Val Val Ile Gly Pro Ser Ser Leu Ile Val His Phe Asn Glu Val
 1070 1075 1080
 Ile Gly Arg Gly His Phe Gly Cys Val Tyr His Gly Thr Leu Leu
 1085 1090 1095
 Asp Asn Asp Gly Lys Lys Ile His Cys Ala Val Lys Ser Leu Asn
 1100 1105 1110
 Arg Ile Thr Asp Ile Gly Glu Val Ser Gln Phe Leu Thr Glu Gly
 1115 1120 1125
 Ile Ile Met Lys Asp Phe Ser His Pro Asn Val Leu Ser Leu Leu
 1130 1135 1140
 Gly Ile Cys Leu Arg Ser Glu Gly Ser Pro Leu Val Val Leu Pro
 1145 1150 1155
 Tyr Met Lys His Gly Asp Leu Arg Asn Phe Ile Arg Asn Glu Thr
 1160 1165 1170
 His Asn Pro Thr Val Lys Asp Leu Ile Gly Phe Gly Leu Gln Val
 1175 1180 1185
 Ala Lys Gly Met Lys Tyr Leu Ala Ser Lys Lys Phe Val His Arg
 1190 1195 1200
 Asp Leu Ala Ala Arg Asn Cys Met Leu Asp Glu Lys Phe Thr Val
 1205 1210 1215

5 Lys Val Ala Asp Phe Gly Leu Ala Arg Asp Met Tyr Asp Lys Glu
 1220 1225 1230
 Tyr Tyr Ser Val His Asn Lys Thr Gly Ala Lys Leu Pro Val Lys
 1235 1240 1245
 10 Trp Met Ala Leu Glu Ser Leu Gln Thr Gln Lys Phe Thr Thr Lys
 1250 1255 1260
 Ser Asp Val Trp Ser Phe Gly Val Val Leu Trp Glu Leu Met Thr
 1265 1270 1275
 15 Arg Gly Ala Pro Pro Tyr Pro Asp Val Asn Thr Phe Asp Ile Thr
 1280 1285 1290
 Val Tyr Leu Leu Gln Gly Arg Arg Leu Leu Gln Pro Glu Tyr Cys
 1295 1300 1305
 20 Pro Asp Pro Leu Tyr Glu Val Met Leu Lys Cys Trp His Pro Lys
 1310 1315 1320
 Ala Glu Met Arg Pro Ser Phe Ser Glu Leu Val Ser Arg Ile Ser
 1325 1330 1335
 25 Ala Ile Phe Ser Thr Phe Ile Gly Glu His Tyr Val His Val Asn
 1340 1345 1350
 30 Ala Thr Tyr Val Asn Val Lys Cys Val Ala Pro Tyr Pro Ser Leu
 1355 1360 1365
 35 Leu Ser Ser Glu Asp Asn Ala Asp Asp Glu Val Asp Thr Arg Pro
 1370 1375 1380
 40 Ala Ser Phe Trp Glu Thr Ser
 1385 1390
 45 <210> 13
 50 <211> 2558
 <212> DNA
 55 <213> Homo sapiens

<220>

5 <221> gene

<222> (1)..(2558)

10 <223> Chondroitin sulfate proteoglycan BEHAB/brevican mRNA, GPI isoform
, complete cds

<220>

15 <221> reference

<222> (1)..(2558)

20 <223> Gary,S.C., Zerillo,C.A., Chiang,V.L., Gaw,J.U., Gray,G. and Hockf
ield,S. TITLE cDNA cloning, chromosomal localization, and express
ion analysis of human BEHAB/brevican, a brain specific proteoglyc
an regulated during cortical development and in glioma JOURNAL Ge
ne 256 (1-2), 139-147 (2000)

<220>

30 <221> CDS

<222> (48)..(2063)

<223>

<400> 13

40 tgtggcactg cctgcgtacc caaccccagc cctgggtagc ctgcagc atg gcc cag 56
Met Ala Gln
1

45 ctg ttc ctg ccc ctg ctg gca gcc ctg gtc ctg gcc cag gct cct gca 104
Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln Ala Pro Ala
5 10 15

50 gct tta gca gat gtt ctg gaa gga gac agc tca gag gac cgc gct ttt 152
Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp Arg Ala Phe
20 25 30 35

55 cgc gtg cgc atc gcg ggc gac gcg cca ctg cag ggc gtg ctc ggc ggc 200
Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val Leu Gly Gly
40 45 50

60 gcc ctc acc atc cct tgc cac gtc cac tac ctg cgg cca ccg ccg agc 248
Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro Pro Pro Ser

55 60 65

5 cgc cgg gct gtg ctg ggc tct ccg cgg gtc aag tgg act ttc ctg tcc 296
Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Trp Thr Phe Leu Ser
70 75 80

10 cgg ggc cgg gag gca gag gtg ctg gtg gcg cgg gga gtg cgc gtc aag 344
Arg Gly Arg Glu Ala Glu Val Leu Val Ala Arg Gly Val Arg Val Lys
85 90 95

15 gtg aac gag gcc tac cgg ttc cgc gtg gca ctg cct gcg tac cca gcg 392
Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala
100 105 110 115

20 tcg ctc acc gac gtc tcc ctg gcg ctg agc gag ctg cgc ccc aac gac 440
Ser Leu Thr Asp Val Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp
120 125 130

25 tca ggt atc tat cgc tgt gag gtc cag cac ggc atc gat gac agc agc 488
Ser Gly Ile Tyr Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser
135 140 145

30 gac gct gtg gag gtc aag gtc aaa ggg gtc gtc ttt ctc tac cga gag 536
Asp Ala Val Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu
150 155 160

35 ggc tct gcc cgc tat gct ttc tcc ttt tct ggg gcc cag gag gcc tgt 584
Gly Ser Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys
165 170 175

40 gcc cgc att gga gcc cac atc gcc acc ccg gag cag ctc tat gcc gcc 632
Ala Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
180 185 190 195

45 tac ctt ggg ggc tat gag caa tgt gat gct ggc tgg ctg tcg gat cag 680
Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp Gln
200 205 210

50 acc gtg agg tat ccc atc cag acc cca cga gag gcc tgt tac gga gac 728
Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr Gly Asp
215 220 225

55 atg gat ggc ttc ccc ggg gtc cgg aac tat ggt gtg gtg gac ccg gat 776
Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val Asp Pro Asp
230 235 240

60 gac ctc tat gat gtg tac tgt tat gct gaa gac cta aat gga gaa ttg 824
Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu Asn Gly Glu Leu
245 250 255

65 ttc ctg ggt gac cct cca gag aag ctg aca ttg gag gaa gca cgg gcg 872
Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu Glu Glu Ala Arg Ala
260 265 270 275

70 tac tgc cag gag cgg ggt gca gag att gcc acc acg ggc caa ctg tat 920
Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala Thr Thr Gly Gln Leu Tyr

280 285 290

5 gca gcc tgg gat ggt ggc ctg gac cac tgc agc cca ggg tgg cta gct 968
Ala Ala Trp Asp Gly Gly Leu Asp His Cys Ser Pro Gly Trp Leu Ala
295 300 305

10 gat ggc agt gtg cgc tac ccc atc gtc aca ccc agc cag cgc tgt ggt 1016
Asp Gly Ser Val Arg Tyr Pro Ile Val Thr Pro Ser Gln Arg Cys Gly
310 315 320

15 ggg ggc ttg cct ggt gtc aag act ctc ttc ctc ttc ccc aac cag act 1064
Gly Gly Leu Pro Gly Val Lys Thr Leu Phe Leu Phe Pro Asn Gln Thr
325 330 335

20 ggc ttc ccc aat aag cac agc cgc ttc aac gtc tac tgc ttc cga gac 1112
Gly Phe Pro Asn Lys His Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp
340 345 350 355

25 tcg gcc cag cct tct gcc atc cct gag gcc tcc aac cca gcc tcc aac 1160
Ser Ala Gln Pro Ser Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn
360 365 370

30 cca gcc tct gat gga cta gag gct atc gtc aca gtg aca gag acc ctg 1208
Pro Ala Ser Asp Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu
375 380 385

35 gag gaa ctg cag ctg cct cag gaa gcc aca gag agt gaa tcc cgt ggg 1256
Glu Glu Leu Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly
390 395 400

40 gcc atc tac tcc atc ccc atc atg gag gac gga gga ggt gga agc tcc 1304
Ala Ile Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser
405 410 415

45 act cca gaa gac cca gca gag gcc cct agg acg ctc cta gaa ttt gaa 1352
Thr Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
420 425 430 435

50 aca caa tcc atg gta ccg ccc acg ggg ttc tca gaa gag gaa ggt aag 1400
Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly Lys
440 445 450

55 gca ttg gag gaa gaa gag aaa tat gaa gat gaa gaa gag aaa gag gag 1448
Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys Glu Glu
455 460 465

60 gaa gaa gaa gag gag gag gtg gag gat gag gct ctg tgg gca tgg ccc 1496
Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp Ala Trp Pro
470 475 480

65 agc gag ctc agc agc ccg ggc cct gag gcc tct ctc ccc act gag cca 1544
Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu Pro Thr Glu Pro
485 490 495

70 gca gcc cag gag gag tca ctc tcc cag gcg cca gca agg gca gtc ctg 1592
Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro Ala Arg Ala Val Leu

	500	505	510	515	
5	cag cct ggt gca tca cca ctt cct gat gga gag tca gaa gct tcc agg Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly Glu Ser Glu Ala Ser Arg	520	525	530	1640
10	cct cca agg gtc cat gga cca cct act gag act ctg ccc act ccc agg Pro Pro Arg Val His Gly Pro Pro Thr Glu Thr Leu Pro Thr Pro Arg	535	540	545	1688
	gag agg aac cta gca tcc cca tca cct tcc act ctg gtt gag gca aga Glu Arg Asn Leu Ala Ser Pro Ser Pro Ser Thr Leu Val Glu Ala Arg	550	555	560	1736
15	gag gtg ggg gag gca act ggt ggt cct gag cta tct ggg gtc cct cga Glu Val Gly Glu Ala Thr Gly Gly Pro Glu Leu Ser Gly Val Pro Arg	565	570	575	1784
20	gga gag agc gag gag aca gga agc tcc gag ggt gcc cct tcc ctg ctt Gly Glu Ser Glu Glu Thr Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu	580	585	590	1832
25	cca gcc aca cgg gcc cct gag ggt acc agg gag ctg gag gcc ccc tct Pro Ala Thr Arg Ala Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser	600	605	610	1880
30	gaa gat aat tct gga aga act gcc cca gca ggg acc tca gtg cag gcc Glu Asp Asn Ser Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala	615	620	625	1928
	cag cca gtg ctg ccc act gac agc gcc agc cga ggt gga gtg gcc gtg Gln Pro Val Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val	630	635	640	1976
35	gtc ccc gca tca ggt aat tct gcc caa ggc tca act gcc ctc tct atc Val Pro Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala Leu Ser Ile	645	650	655	2024
40	cta ctc ctt ttc ttc ccc ctg cag ctc tgg gtc acc tga cctgtagtcc Leu Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr	660	665	670	2073
	tttaaccac catcatccca aactctctg tcctttgcct tcattctctt acccacctct				2133
45	acctatgggt ctccaatctc ggatateccac cttgtgggta tctcagctct ccgcgtcttt				2193
	accctgtgat ccagccccg ccaactgacca tctgtgaccc ttcctgcca ttgggccctc				2253
50	cacctgtggc tcacatctcg ccagccccac agagcatcct caggcctctc caagggctct				2313
	catcacctat tgcagccttc agggctcggc ctattttcca ctactccctt catccgcctg				2373
	tgtgccgtcc ccttttagctg cctcctattg atctcagga agcctgggag tcccttctca				2433
55	cccctcaacc tccggagtcc aggagaacct gtacccccac agagccttaa gcaactactt				2493

ctgtgaagta ttttttgact gtttcatgga aaacaagcct tggaaataaa tctctattaa 2553
accgc 2558

5

<210> 14
<211> 671

10

<212> PRT
<213> Homo sapiens

15

<220>
<221> gene
<222> (1) .. (671)
<223> Chondroitin sulfate proteoglycan BEHAB/brevican

20

<400> 14
Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln
1 5 10 15

25

Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp
20 25 30

30

Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val
35 40 45

35

Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro
50 55 60

40

Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Trp Thr
65 70 75 80

45

Phe Leu Ser Arg Gly Arg Glu Ala Glu Val Leu Val Ala Arg Gly Val
85 90 95

50

Arg Val Lys Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala
100 105 110

55

Tyr Pro Ala Ser Leu Thr Asp Val Ser Leu Ala Leu Ser Glu Leu Arg

115 120 125

5 Pro Asn Asp Ser Gly Ile Tyr Arg Cys Glu Val Gln His Gly Ile Asp
130 135 140

10 Asp Ser Ser Asp Ala Val Glu Val Lys Val Lys Gly Val Val Phe Leu
145 150 155 160

15 Tyr Arg Glu Gly Ser Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln
165 170 175

20 Glu Ala Cys Ala Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu
180 185 190

25 Tyr Ala Ala Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu
195 200 205

30 Ser Asp Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys
210 215 220

35 Tyr Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
225 230 235 240

40 Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu Asn
245 250 255

45 Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu Glu Glu
260 265 270

50 Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala Thr Thr Gly
275 280 285

55 Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His Cys Ser Pro Gly
290 295 300

60 Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile Val Thr Pro Ser Gln
305 310 315 320

65 Arg Cys Gly Gly Gly Leu Pro Gly Val Lys Thr Leu Phe Leu Phe Pro
325 330 335

70 Asn Gln Thr Gly Phe Pro Asn Lys His Ser Arg Phe Asn Val Tyr Cys

340 345 350

5 Phe Arg Asp Ser Ala Gln Pro Ser Ala Ile Pro Glu Ala Ser Asn Pro
355 360 365

10 Ala Ser Asn Pro Ala Ser Asp Gly Leu Glu Ala Ile Val Thr Val Thr
370 375 380

15 Glu Thr Leu Glu Glu Leu Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu
385 390 395 400

20 Ser Arg Gly Ala Ile Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly
405 410 415

25 Gly Ser Ser Thr Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu
420 425 430

30 Glu Phe Glu Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu
435 440 445

35 Glu Gly Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu
450 455 460

40 Lys Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
465 470 475 480

45 Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu Pro
485 490 495

50 Thr Glu Pro Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro Ala Arg
500 505 510

55 Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly Glu Ser Glu
515 520 525

60 Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr Glu Thr Leu Pro
530 535 540

65 Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser Pro Ser Thr Leu Val
545 550 555 560

70 Glu Ala Arg Glu Val Gly Glu Ala Thr Gly Gly Pro Glu Leu Ser Gly

565

570

575

5

Val Pro Arg Gly Glu Ser Glu Glu Thr Gly Ser Ser Glu Gly Ala Pro
580 585 590

10

Ser Leu Leu Pro Ala Thr Arg Ala Pro Glu Gly Thr Arg Glu Leu Glu
595 600 605

15

Ala Pro Ser Glu Asp Asn Ser Gly Arg Thr Ala Pro Ala Gly Thr Ser
610 615 620

Val Gln Ala Gln Pro Val Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly
625 630 635 640

20

Val Ala Val Val Pro Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala
645 650 655

25

Leu Ser Ile Leu Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr
660 665 670

30

<210> 15

<211> 2316

<212> DNA

35

<213> Homo sapiens

<220>

40

<221> gene

<222> (1) .. (2316)

45

<223> Human mRNA for CD44E (epithelial form)

<220>

50

<221> REFERENCE

<222> (1) .. (2316)

55

<223> Stamenkovic, I., Aruffo, A., Amiot, M. and Seed, B. TITLE The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells JOU

RNAL EMBO J. 10 (2), 343-348 (1991)

5 <220>
 <221> CDS
 <222> (124)..(1605)
 10 <223>
 15 <400> 15
 agcggacccc agcctctgcc aggttcggtc cgccatcctc gtcccgtcct ccgcccggccc 60
 ctgccccgcg cccagggatc ctccagctcc ttctgcccgc gccctccgtt cgctccggac 120
 20 acc atg gac aag ttt tgg tgg cac gca gcc tgg gga ctc tgc ctc gtg 168
 Met Asp Lys Phe Trp Trp His Ala Ala Trp Gly Leu Cys Leu Val
 1 5 10 15
 ccg ctg agc ctg gcg cag atc gat ttg aat ata acc tgc cgc ttt gca 216
 25 Pro Leu Ser Leu Ala Gln Ile Asp Leu Asn Ile Thr Cys Arg Phe Ala
 20 25 30
 ggt gta ttc cac gtg gag aaa aat ggt cgc tac agc atc tct cgg acg 264
 Gly Val Phe His Val Glu Lys Asn Gly Arg Tyr Ser Ile Ser Arg Thr
 30 35 40 45
 gag gcc gct gac ctc tgc aag gct ttc aat agc acc ttg ccc aca atg 312
 Glu Ala Ala Asp Leu Cys Lys Ala Phe Asn Ser Thr Leu Pro Thr Met
 50 55 60
 35 gcc cag atg gag aaa gct ctg agc atc gga ttt gag acc tgc agg tat 360
 Ala Gln Met Glu Lys Ala Leu Ser Ile Gly Phe Glu Thr Cys Arg Tyr
 65 70 75
 40 ggg ttc ata gaa ggg cat gtg gtg att ccc cgg atc cac ccc aac tcc 408
 Gly Phe Ile Glu Gly His Val Val Ile Pro Arg Ile His Pro Asn Ser
 80 85 90 95
 45 atc tgt gca gca aac aac aca ggg gtg tac atc ctc aca tac aac acc 456
 Ile Cys Ala Ala Asn Asn Thr Gly Val Tyr Ile Leu Thr Tyr Asn Thr
 100 105 110
 tcc cag tat gac aca tat tgc ttc aat gct tca gct cca cct gaa gaa 504
 50 Ser Gln Tyr Asp Thr Tyr Cys Phe Asn Ala Ser Ala Pro Pro Glu Glu
 115 120 125
 gat tgt aca tca gtc aca gac ctg ccc aat gcc ttt gat gga cca att 552
 Asp Cys Thr Ser Val Thr Asp Leu Pro Asn Ala Phe Asp Gly Pro Ile
 130 135 140
 55 acc ata act att gtt aac cgt gat ggc acc cgc tat gtc cag aaa gga 600

Thr Ile Thr Ile Val Asn Arg Asp Gly Thr Arg Tyr Val Gln Lys Gly
 145 150 155
 5 gaa tac aga acg aat cct gaa gac atc tac ccc agc aac cct act gat 648
 Glu Tyr Arg Thr Asn Pro Glu Asp Ile Tyr Pro Ser Asn Pro Thr Asp
 160 165 170 175
 10 gat gac gtg agc agc ggc tcc tcc agt gaa agg agc agc act tca gga 696
 Asp Asp Val Ser Ser Gly Ser Ser Ser Glu Arg Ser Ser Thr Ser Gly
 180 185 190
 15 ggt tac atc ttt tac acc ttt tct act gta cac ccc atc cca gac gaa 744
 Gly Tyr Ile Phe Tyr Thr Phe Ser Thr Val His Pro Ile Pro Asp Glu
 195 200 205
 gac agt ccc tgg atc acc gac agc aca gac aga atc cct cgt acc aat 792
 Asp Ser Pro Trp Ile Thr Asp Ser Thr Asp Arg Ile Pro Arg Thr Asn
 210 215 220
 20 atg gac tcc agt cat agt aca acg ctt cag cct act gca aat cca aac 840
 Met Asp Ser Ser His Ser Thr Thr Leu Gln Pro Thr Ala Asn Pro Asn
 225 230 235
 25 aca ggt ttg gtg gaa gat ttg gac agg aca gga cct ctt tca atg aca 888
 Thr Gly Leu Val Glu Asp Leu Asp Arg Thr Gly Pro Leu Ser Met Thr
 240 245 250 255
 30 acg cag cag agt aat tct cag agc ttc tct aca tca cat gaa ggc ttg 936
 Thr Gln Gln Ser Asn Ser Gln Ser Phe Ser Thr Ser His Glu Gly Leu
 260 265 270
 gaa gaa gat aaa gac cat cca aca act tct act ctg aca tca agc aat 984
 Glu Glu Asp Lys Asp His Pro Thr Thr Ser Thr Leu Thr Ser Ser Asn
 275 280 285
 35 agg aat gat gtc aca ggt gga aga aga gac cca aat cat tct gaa ggc 1032
 Arg Asn Asp Val Thr Gly Gly Arg Arg Asp Pro Asn His Ser Glu Gly
 290 295 300
 40 tca act cat tta ctg gaa ggt tat acc tct cat tac cca cac acg aag 1080
 Ser Thr His Leu Leu Glu Gly Tyr Thr Ser His Tyr Pro His Thr Lys
 305 310 315
 45 gaa agc agg acc ttc atc cca gtg acc tca gct aag act ggg tcc ttt 1128
 Glu Ser Arg Thr Phe Ile Pro Val Thr Ser Ala Lys Thr Gly Ser Phe
 320 325 330 335
 50 gga gtt act gca gtt act gtt gga gat tcc aac tct aat gtc aat cgt 1176
 Gly Val Thr Ala Val Thr Val Gly Asp Ser Asn Ser Asn Val Asn Arg
 340 345 350
 55 tcc tta tca gga gac caa gac aca ttc cac ccc agt ggg ggg tcc cat 1224
 Ser Leu Ser Gly Asp Gln Asp Thr Phe His Pro Ser Gly Gly Ser His
 355 360 365
 acc act cat gga tct gaa tca gat gga cac tca cat ggg agt caa gaa 1272

Thr Thr His Gly Ser Glu Ser Asp Gly His Ser His Gly Ser Gln Glu
370 375 380

5 ggt gga gca aac aca acc tct ggt cct ata agg aca ccc caa att cca 1320
Gly Gly Ala Asn Thr Thr Ser Gly Pro Ile Arg Thr Pro Gln Ile Pro
385 390 395

10 gaa tgg ctg atc atc ttg gca tcc ctc ttg gcc ttg gct ttg att ctt 1368
Glu Trp Leu Ile Ile Leu Ala Ser Leu Leu Ala Leu Ala Leu Ile Leu
400 405 410 415

15 gca gtt tgc att gca gtc aac agt cga aga agg tgt ggg cag aag aaa 1416
Ala Val Cys Ile Ala Val Asn Ser Arg Arg Arg Cys Gly Gln Lys Lys
420 425 430

20 aag cta gtg atc aac agt ggc aat gga gct gtg gag gac aga aag cca 1464
Lys Leu Val Ile Asn Ser Gly Asn Gly Ala Val Glu Asp Arg Lys Pro
435 440 445

25 agt gga ctc aac gga gag gcc agc aag tct cag gaa atg gtg cat ttg 1512
Ser Gly Leu Asn Gly Glu Ala Ser Lys Ser Gln Glu Met Val His Leu
450 455 460

30 gtg aac aag gag tcg tca gaa act cca gac cag ttt atg aca gct gat 1560
Val Asn Lys Glu Ser Ser Glu Thr Pro Asp Gln Phe Met Thr Ala Asp
465 470 475

35 gag aca agg aac ctg cag aat gtg gac atg aag att ggg gtg taa 1605
Glu Thr Arg Asn Leu Gln Asn Val Asp Met Lys Ile Gly Val
480 485 490

40 cacctacacc attatcttgg aaagaaacaa cgttggaaac ataaccatta caggggagct 1665

45 gggacactta acagatgcaa tgtgctactg attgtttcat ttcgaatcta taatagcata 1725

50 aaattttcta ctctttttgt tttttgtgtt ttgttcttta aagtcagggtc caatttgtaa 1785

55 aaacagcatt gctttctgaa attagggccc aattaataat cagcaagaat tttgatcggt 1845

60 tcagttcccc acttgagggc ctttcatccc tcgggtgtgc tatggatggc ttctaacaaa 1905

65 aacctaccac atagttattc ctgatcgcca accttgcccc ccaccagcta aggacatttc 1965

70 caggggtaat agggcctggt cctgggagga aatttgaatg ggtcattttg cccttccatt 2025

75 agcctaatacc ctgggcattg ctttccactg aggttggggg ttgggggtgta ctagttacac 2085

80 atcttcaaca gacccccctc agaaatTTTT cagatgcttc tgggagacac ccaaagggta 2145

85 agtctatttta tctgtagtaa actatTTTatc tgtgtTTTTg aaatattaaa ccctggatca 2205

90 gtccttttat tcagtataat tttttaagT tactttgtca gaggcacaaa aagggtttta 2265

95 actgattcat aataaatatc tgtaccttct tcgaaaaaaa aaaaaaaaaa a 2316

5 <210> 16
<211> 742
<212> PRT
<213> Homo sapiens
10 <220>
<221> SIGNAL
15 <222> (1) .. (20)
<223> BY SIMILARITY
20 <220>
<221> Gene
25 <222> (1) .. (742)
<223> CD-44 Antigen
30 <220>
<221> CHAIN
35 <222> (21) .. (742)
<223> CD44 ANTIGEN
40 <220>
<221> DOMAIN
45 <222> (21) .. (649)
<223> EXTRACELLULAR (POTENTIAL)
50 <220>
<221> TRANSMEM
55 <222> (650) .. (670)

<223> POTENTIAL

5 <220>

<221> DOMAIN

<222> (671)..(742)

10

<223> CYTOPLASMIC (POTENTIAL)

15 <220>

<221> DOMAIN

<222> (48)..(119)

20

<223> LINK

25 <220>

<221> MOD_RES

<222> (21)..(21)

30

<223> PYRROLIDONE CARBOXYLIC ACID (PROBABLE)

35 <220>

<221> DISULFID

<222> (53)..(118)

40

<223> BY SIMILARITY

45 <220>

<221> DISULFID

<222> (77)..(97)

50

<223> BY SIMILARITY

55 <220>

<221> CARBOHYD

<222> (25) .. (25)

5 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

10

<221> CARBOHYD

<222> (57) .. (57)

15 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

20

<221> CARBOHYD

<222> (100) .. (100)

25 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

30

<221> CARBOHYD

<222> (110) .. (110)

35 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

40

<221> CARBOHYD

<222> (120) .. (120)

45 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

50

<221> CARBOHYD

<222> (350) .. (350)

55 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

5 <221> CARBOHYD

<222> (548)..(548)

10 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

15 <221> CARBOHYD

<222> (599)..(599)

20 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

25 <221> CARBOHYD

<222> (636)..(636)

30 <223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>

35 <221> VARSPLIC

<222> (23)..(29)

40 <223> DLNITCR -> GVGRRKS (IN ISOFORM CD44SP)

<220>

45 <221> VARSPLIC

<222> (30)..(742)

50 <223> MISSING (IN ISOFORM CD44SP)

<220>

55 <221> VARSPLIC

<222> (192)..(192)

<223> G -> A (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXON
5)

5

<220>

10 <221> VARSPLIC

<222> (193)..(223)

15 <223> MISSING (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXO
N 5)

<220>

<221> VARSPLIC

<222> (223)..(223)

<223> T -> S (IN ISOFORM WITHOUT EXON 6)

<220>

<221> VARSPLIC

<222> (224)..(266)

<223> MISSING (IN ISOFORM WITHOUT EXON 6)

<220>

<221> VARSPLIC

<222> (223)..(223)

<223> T -> N (IN ISOFORMS WITHOUT EXONS 6 TO 11)

<220>

<221> VARSPLIC

<222> (224)..(472)

<223> MISSING (IN ISOFORMS WITHOUT EXONS 6 TO 11)

<220>

5 <221> VARSPLIC

<222> (223)..(223)

10 <223> T -> R (IN ISOFORMS WITHOUT EXONS 6 TO 14)

<220>

15 <221> VARSPLIC

<222> (224)..(604)

<223> MISSING (IN ISOFORMS WITHOUT EXONS 6 TO 14)

<220>

25 <221> VARSPLIC

<222> (266)..(273)

<223> MISSING (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXO
N 7)

<220>

35 <221> VARSPLIC

<222> (385)..(385)

40 <223> I -> T (IN ISOFORMS WITHOUT EXON 10)

<220>

45 <221> VARSPLIC

<222> (386)..(428)

50 <223> MISSING (IN ISOFORMS WITHOUT EXON 10)

<220>

55 <221> VARSPLIC

<222> (506)..(506)

<223> Q -> R (IN ISOFORMS WITHOUT EXON 13)

<220>

<221> VARSPLIC

<222> (507)..(535)

<223> MISSING (IN ISOFORMS WITHOUT EXON 13)

<220>

<221> VARSPLIC

<222> (536)..(536)

<223> N -> R (IN ISOFORMS WITHOUT EXON 14)

<220>

<221> VARSPLIC

<222> (537)..(604)

<223> MISSING (IN ISOFORMS WITHOUT EXON 14)

<220>

<221> VARSPLIC

<222> (675)..(675)

<223> R -> S (IN ISOFORMS WITHOUT EXON 19)

<220>

<221> VARSPLIC

<222> (676)..(742)

<223> MISSING (IN ISOFORMS WITHOUT EXON 19)

<220>

<221> VARIANT

<222> (46) .. (46)

<223> R -> P (IN IN(A) ANTIGEN)

<400> 16

Met Asp Lys Phe Trp Trp His Ala Ala Trp Gly Leu Cys Leu Val Pro
1 5 10 15

Leu Ser Leu Ala Gln Ile Asp Leu Asn Ile Thr Cys Arg Phe Ala Gly
20 25 30

Val Phe His Val Glu Lys Asn Gly Arg Tyr Ser Ile Ser Arg Thr Glu
35 40 45

Ala Ala Asp Leu Cys Lys Ala Phe Asn Ser Thr Leu Pro Thr Met Ala
50 55 60

Gln Met Glu Lys Ala Leu Ser Ile Gly Phe Glu Thr Cys Arg Tyr Gly
65 70 75 80

Phe Ile Glu Gly His Val Val Ile Pro Arg Ile His Pro Asn Ser Ile
85 90 95

Cys Ala Ala Asn Asn Thr Gly Val Tyr Ile Leu Thr Ser Asn Thr Ser
100 105 110

Gln Tyr Asp Thr Tyr Cys Phe Asn Ala Ser Ala Pro Pro Glu Glu Asp
115 120 125

Cys Thr Ser Val Thr Asp Leu Pro Asn Ala Phe Asp Gly Pro Ile Thr
130 135 140

Ile Thr Ile Val Asn Arg Asp Gly Thr Arg Tyr Val Gln Lys Gly Glu
145 150 155 160

Tyr Arg Thr Asn Pro Glu Asp Ile Tyr Pro Ser Asn Pro Thr Asp Asp
165 170 175

Asp Val Ser Ser Gly Ser Ser Ser Glu Arg Ser Ser Thr Ser Gly Gly
180 185 190

5
Tyr Ile Phe Tyr Thr Phe Ser Thr Val His Pro Ile Pro Asp Glu Asp
195 200 205

10
Ser Pro Trp Ile Thr Asp Ser Thr Asp Arg Ile Pro Ala Thr Thr Leu
210 215 220

15
Met Ser Thr Ser Ala Thr Ala Thr Glu Thr Ala Thr Lys Arg Gln Glu
225 230 235 240

20
Thr Trp Asp Trp Phe Ser Trp Leu Phe Leu Pro Ser Glu Ser Lys Asn
245 250 255

25
His Leu His Thr Thr Thr Gln Met Ala Gly Thr Ser Ser Asn Thr Ile
260 265 270

30
Ser Ala Gly Trp Glu Pro Asn Glu Glu Asn Glu Asp Glu Arg Asp Arg
275 280 285

35
His Leu Ser Phe Ser Gly Ser Gly Ile Asp Asp Asp Glu Asp Phe Ile
290 295 300

40
Ser Ser Thr Ile Ser Thr Thr Pro Arg Ala Phe Asp His Thr Lys Gln
305 310 315 320

45
Asn Gln Asp Trp Thr Gln Trp Asn Pro Ser His Ser Asn Pro Glu Val
325 330 335

50
Leu Leu Gln Thr Thr Thr Arg Met Thr Asp Val Asp Arg Asn Gly Thr
340 345 350

55
Thr Ala Tyr Glu Gly Asn Trp Asn Pro Glu Ala His Pro Pro Leu Ile
355 360 365

60
His His Glu His His Glu Glu Glu Glu Thr Pro His Ser Thr Ser Thr
370 375 380

65
Ile Gln Ala Thr Pro Ser Ser Thr Thr Glu Glu Thr Ala Thr Gln Lys
385 390 395 400

5
10
15
20
25
30
35
40
45
50
55

Glu Gln Trp Phe Gly Asn Arg Trp His Glu Gly Tyr Arg Gln Thr Pro
405 410 415

Arg Glu Asp Ser His Ser Thr Thr Gly Thr Ala Ala Ala Ser Ala His
420 425 430

Thr Ser His Pro Met Gln Gly Arg Thr Thr Pro Ser Pro Glu Asp Ser
435 440 445

Ser Trp Thr Asp Phe Phe Asn Pro Ile Ser His Pro Met Gly Arg Gly
450 455 460

His Gln Ala Gly Arg Arg Met Asp Met Asp Ser Ser His Ser Thr Thr
465 470 475 480

Leu Gln Pro Thr Ala Asn Pro Asn Thr Gly Leu Val Glu Asp Leu Asp
485 490 495

Arg Thr Gly Pro Leu Ser Met Thr Thr Gln Gln Ser Asn Ser Gln Ser
500 505 510

Phe Ser Thr Ser His Glu Gly Leu Glu Glu Asp Lys Asp His Pro Thr
515 520 525

Thr Ser Thr Leu Thr Ser Ser Asn Arg Asn Asp Val Thr Gly Gly Arg
530 535 540

Arg Asp Pro Asn His Ser Glu Gly Ser Thr Thr Leu Leu Glu Gly Tyr
545 550 555 560

Thr Ser His Tyr Pro His Thr Lys Glu Ser Arg Thr Phe Ile Pro Val
565 570 575

Thr Ser Ala Lys Thr Gly Ser Phe Gly Val Thr Ala Val Thr Val Gly
580 585 590

Asp Ser Asn Ser Asn Val Asn Arg Ser Leu Ser Gly Asp Gln Asp Thr
595 600 605

Phe His Pro Ser Gly Gly Ser His Thr Thr His Gly Ser Glu Ser Asp
610 615 620

Gly His Ser His Gly Ser Gln Glu Gly Gly Ala Asn Thr Thr Ser Gly
625 630 635 640

5

Pro Ile Arg Thr Pro Gln Ile Pro Glu Trp Leu Ile Ile Leu Ala Ser
645 650 655

10

Leu Leu Ala Leu Ala Leu Ile Leu Ala Val Cys Ile Ala Val Asn Ser
660 665 670

15

Arg Arg Arg Cys Gly Gln Lys Lys Lys Leu Val Ile Asn Ser Gly Asn
675 680 685

20

Gly Ala Val Glu Asp Arg Lys Pro Ser Gly Leu Asn Gly Glu Ala Ser
690 695 700

25

Lys Ser Gln Glu Met Val His Leu Val Asn Lys Glu Ser Ser Glu Thr
705 710 715 720

30

Pro Asp Gln Phe Met Thr Ala Asp Glu Thr Arg Asn Leu Gln Asn Val
725 730 735

35

Asp Met Lys Ile Gly Val
740

<210> 17

<211> 930

<212> DNA

40

<213> Homo sapiens

45

<220>

<221> REFERENCE

<222> (1)..(930)

50

<223> Todd, S.C., Doctor, V.S. and Levy, S. TITLE Sequences and express
ion of six new members of the tetraspanin/TM4SF family JOURNAL Bi
ochim. Biophys. Acta 1399 (1), 101-104 (1998)

55

<220>

<221> Gene

<222> (1)..(930)

<223> Tetraspan TM4SF (TSPAN-3), mRNA

<220>

<221> CDS

<222> (1)..(762)

<223>

<400> 17

atg ggc cag tgc ggc atc acc tcc tcc aag acc gtg ctg gtc ttt ctc	48
Met Gly Gln Cys Gly Ile Thr Ser Ser Lys Thr Val Leu Val Phe Leu	
1 5 10 15	

aac ctc atc ttc tgg ggg gca gct ggc att tta tgc tat gtg gga gcc	96
Asn Leu Ile Phe Trp Gly Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala	
20 25 30	

tat gtc ttc atc act tat gat gac tat gac cac ttc ttt gaa gat gtg	144
Tyr Val Phe Ile Thr Tyr Asp Asp Tyr Asp His Phe Phe Glu Asp Val	
35 40 45	

tac acg ctc atc cct gct gta gtg atc ata gct gta gga gcc ctg ctt	192
Tyr Thr Leu Ile Pro Ala Val Val Ile Ile Ala Val Gly Ala Leu Leu	
50 55 60	

ttc atc att ggg cta att ggc tgc tgt gcc aca atc cgg gaa agt cgc	240
Phe Ile Ile Gly Leu Ile Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg	
65 70 75 80	

tgt gga ctt gcc acg ttt gtc atc atc ctg ctc ttg gtt ttt gtc aca	288
Cys Gly Leu Ala Thr Phe Val Ile Ile Leu Leu Leu Val Phe Val Thr	
85 90 95	

gaa gtt gtt gta gtg gtt ttg gga tat gtt tac aga gca aag gtg gaa	336
Glu Val Val Val Val Val Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu	
100 105 110	

aat gag gtt gat cgc agc att cag aaa gtg tat aag acc tac aat gga	384
Asn Glu Val Asp Arg Ser Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly	
115 120 125	

acc aac cct gat gct gct agc cgg gct att gat tat gta cag aga cag	432
Thr Asn Pro Asp Ala Ala Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln	
130 135 140	

ctg cat tgt tgt gga att cac aac tac tca gac tgg gaa aat aca gat 480
Leu His Cys Cys Gly Ile His Asn Tyr Ser Asp Trp Glu Asn Thr Asp
145 150 155 160

5 tgg ttc aaa gaa acc aaa aac cag agt gtc cct ctt agc tgc tgc aga 528
Trp Phe Lys Glu Thr Lys Asn Gln Ser Val Pro Leu Ser Cys Cys Arg
165 170 175

10 gag act gcc agc aat tgt aat ggc agc ctg gcc cac cct tcc gac ctc 576
Glu Thr Ala Ser Asn Cys Asn Gly Ser Leu Ala His Pro Ser Asp Leu
180 185 190

15 tat gct gag ggg tgt gag gct cta gtt gtg aag aag cta caa gaa atc 624
Tyr Ala Glu Gly Cys Glu Ala Leu Val Val Lys Lys Leu Gln Glu Ile
195 200 205

20 atg atg cat gtg atc tgg gcc gca ctg gca ttt gca gct att cag ctg 672
Met Met His Val Ile Trp Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu
210 215 220

25 ctg ggc atg ctg tgt gct tgc atc gtg ttg tgc aga agg agt aga gat 720
Leu Gly Met Leu Cys Ala Cys Ile Val Leu Cys Arg Arg Ser Arg Asp
225 230 235 240

30 cct gct tac gag ctc ctc atc act ggc gga acc tat gca tag 762
Pro Ala Tyr Glu Leu Leu Ile Thr Gly Gly Thr Tyr Ala
245 250

35 ttgacaactc ttgacctgagc tttttggtct tgttctgatt tggaaggtga attgagcagg 822
tctgctgctg ttggcctctg gagttcattt agttaagca catgtacact ggtgttggac 882
agagcagctt ggcttttcat gtgcccaact acttactact actgcat 930

40 <210> 18
<211> 253
<212> PRT
<213> Homo sapiens

45 <220>
<221> DOMAIN

50 <222> (1)..(11)
<223> Cytoplasmic (Potential)

55 <220>

<221> Gene

<222> (1) .. (253)

<223> Tetraspanin-3

<220>

<221> TRANSMEM

<222> (12) .. (32)

<223> Potential

<220>

<221> DOMAIN

<222> (33) .. (50)

<223> Extracellular (Potential)

<220>

<221> TRANSMEM

<222> (51) .. (71)

<223> Potential

<220>

<221> DOMAIN

<222> (72) .. (85)

<223> Cytoplasmic (potential)

<220>

<221> TRANSMEM

<222> (86) .. (106)

<223> Potential

5 <220>
<221> DOMAIN
<222> (107)..(212)
10 <223> Extracellular (potential)

15 <220>
<221> TRANSMEM
<222> (213)..(233)
20 <223> Potential

25 <220>
<221> DOMAIN
<222> (234)..(253)
30 <223> Cytoplasmic (potential)

35 <220>
<221> CARBOHYD
<222> (127)..(127)
40 <223> N-linked (GLCNAC...) (potential)

45 <220>
<221> CARBOHYD
<222> (152)..(152)
50 <223> N-linked (GLCNAC...) (potential)

55 <220>
<221> CARBOHYD

<222> (167)..(167)

<223> N-linked (GLCNAC...) (potential)

<220>

<221> CARBOHYD

<222> (183)..(183)

<223> N-linked (GLCNAC...) (potential)

<400> 18

Met Gly Gln Cys Gly Ile Thr Ser Ser Lys Thr Val Leu Val Phe Leu
1 5 10 15

Asn Leu Ile Phe Trp Gly Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala
20 25 30

Tyr Val Phe Ile Thr Tyr Asp Asp Tyr Asp His Phe Phe Glu Asp Val
35 40 45

Tyr Thr Leu Ile Pro Ala Val Val Ile Ile Ala Val Gly Ala Leu Leu
50 55 60

Phe Ile Ile Gly Leu Ile Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg
65 70 75 80

Cys Gly Leu Ala Thr Phe Val Ile Ile Leu Leu Leu Val Phe Val Thr
85 90 95

Glu Val Val Val Val Val Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu
100 105 110

Asn Glu Val Asp Arg Ser Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly
115 120 125

Thr Asn Pro Asp Ala Ala Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln
130 135 140

Leu His Cys Cys Gly Ile His Asn Tyr Ser Asp Trp Glu Asn Thr Asp

145 150 155 160

5 Trp Phe Lys Glu Thr Lys Asn Gln Ser Val Pro Leu Ser Cys Cys Arg
165 170 175

10 Glu Thr Ala Ser Asn Cys Asn Gly Ser Leu Ala His Pro Ser Asp Leu
180 185 190

Tyr Ala Glu Gly Cys Glu Ala Leu Val Val Lys Lys Leu Gln Glu Ile
195 200 205

15 Met Met His Val Ile Trp Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu
210 215 220

20 Leu Gly Met Leu Cys Ala Cys Ile Val Leu Cys Arg Arg Ser Arg Asp
225 230 235 240

25 Pro Ala Tyr Glu Leu Leu Ile Thr Gly Gly Thr Tyr Ala
245 250

30 <210> 19

<211> 1317

<212> DNA

35 <213> Homo sapiens

<220>

40 <221> REFERENCE

<222> (1) .. (1317)

45 <223> Svoboda, M., Tastenoy, M., Van Rampelbergh, J., Goossens, J.F., D
e Neef, P., Waelbroeck, M. and Robberecht, P. TITLE Molecular clo
ning and functional characterization of a human VIP receptor from
SUP-T1 lymphoblasts JOURNAL Biochem. Biophys. Res. Commun. 205 (3), 1617-1624 (1994)

50 <220>

55 <221> Gene

<222> (1) .. (1317)

<223> Vasoactive Intestinal Peptide Receptor-2

5

<220>

<221> CDS

10

<222> (1) .. (1317)

<223>

15

<400> 19

atg	cgg	acg	ctg	ctg	cct	ccc	gcg	ctg	ctg	acc	tgc	tgg	ctg	ctc	gcc	48
Met	Arg	Thr	Leu	Leu	Pro	Pro	Ala	Leu	Leu	Thr	Cys	Trp	Leu	Leu	Ala	
1			5					10					15			

ccc	gtg	aac	agc	att	cac	cca	gaa	tgc	cga	ttt	cat	ctg	gaa	ata	cag	96
Pro	Val	Asn	Ser	Ile	His	Pro	Glu	Cys	Arg	Phe	His	Leu	Glu	Ile	Gln	
		20					25					30				

gag	gaa	gaa	aca	aaa	tgt	aca	gag	ctt	ctg	agg	tct	caa	aca	gaa	aaa	144
Glu	Glu	Glu	Thr	Lys	Cys	Thr	Glu	Leu	Leu	Arg	Ser	Gln	Thr	Glu	Lys	
		35					40					45				

cac	aaa	gcc	tgc	agt	ggc	gtc	tgg	gac	aac	atc	acg	tgc	tgg	cgg	cct	192
His	Lys	Ala	Cys	Ser	Gly	Val	Trp	Asp	Asn	Ile	Thr	Cys	Trp	Arg	Pro	
	50					55					60					

gcc	aat	gtg	gga	gag	acc	gtc	acg	gtg	ccc	tgc	cca	aaa	gtc	ttc	agc	240
Ala	Asn	Val	Gly	Glu	Thr	Val	Thr	Val	Pro	Cys	Pro	Lys	Val	Phe	Ser	
65					70				75					80		

aat	ttt	tac	agc	aaa	gca	gga	aac	ata	agc	aaa	aac	tgt	acg	agt	gac	288
Asn	Phe	Tyr	Ser	Lys	Ala	Gly	Asn	Ile	Ser	Lys	Asn	Cys	Thr	Ser	Asp	
			85					90						95		

gga	tgg	tca	gag	acg	ttc	cca	gat	ttc	gtc	gat	gcc	tgt	ggc	tac	agc	336
Gly	Trp	Ser	Glu	Thr	Phe	Pro	Asp	Phe	Val	Asp	Ala	Cys	Gly	Tyr	Ser	
		100						105					110			

gac	cgg	gag	gat	gag	agc	aag	atc	acg	ttt	tat	att	ctg	gtg	aag	gcc	384
Asp	Pro	Glu	Asp	Glu	Ser	Lys	Ile	Thr	Phe	Tyr	Ile	Leu	Val	Lys	Ala	
		115					120					125				

att	tat	acc	ctg	ggc	tac	agt	gtc	tct	ctg	atg	tct	ctt	gca	aca	gga	432
Ile	Tyr	Thr	Leu	Gly	Tyr	Ser	Val	Ser	Leu	Met	Ser	Leu	Ala	Thr	Gly	
	130					135					140					

agc	ata	att	ctg	tgc	ctc	ttc	agg	aag	ctg	cac	tgc	acc	agg	aat	tac	480
Ser	Ile	Ile	Leu	Cys	Leu	Phe	Arg	Lys	Leu	His	Cys	Thr	Arg	Asn	Tyr	
145					150					155					160	

	atc cac ctg aac ctg ttc ctg tcc ttc atc ctg aga gcc atc tca gtg	528
	Ile His Leu Asn Leu Phe Leu Ser Phe Ile Leu Arg Ala Ile Ser Val	
	165 170 175	
5	ctg gtc aag gac gac gtt ctc tac tcc agc tct ggc acg ttg cac tgc	576
	Leu Val Lys Asp Asp Val Leu Tyr Ser Ser Ser Gly Thr Leu His Cys	
	180 185 190	
10	cct gac cag cca tcc tcc tgg gtg ggc tgc aag ctg agc ctg gtc ttc	624
	Pro Asp Gln Pro Ser Ser Trp Val Gly Cys Lys Leu Ser Leu Val Phe	
	195 200 205	
15	ctg cag tac tgc atc atg gcc aac ttc ttc tgg ctg ctg gtg gag ggg	672
	Leu Gln Tyr Cys Ile Met Ala Asn Phe Phe Trp Leu Leu Val Glu Gly	
	210 215 220	
20	ctc tac ctc cac acc ctc ctg gtg gcc atg ctc ccc cct aga agg tgc	720
	Leu Tyr Leu His Thr Leu Leu Val Ala Met Leu Pro Pro Arg Arg Cys	
	225 230 235 240	
25	ttc ctg gcc tac ctc ctg atc gga tgg ggc ctc ccc acc gtc tgc atc	768
	Phe Leu Ala Tyr Leu Leu Ile Gly Trp Gly Leu Pro Thr Val Cys Ile	
	245 250 255	
30	ggt gca tgg act gcg gcc agg ctc tac tta gaa gac acc ggt tgc tgg	816
	Gly Ala Trp Thr Ala Ala Arg Leu Tyr Leu Glu Asp Thr Gly Cys Trp	
	260 265 270	
35	gat aca aac gac cac agt gtg ccc tgg tgg gtc ata cga ata ccg att	864
	Asp Thr Asn Asp His Ser Val Pro Trp Trp Val Ile Arg Ile Pro Ile	
	275 280 285	
40	tta att tcc atc atc gtc aat ttt gtc ctt ttc att agt att ata cga	912
	Leu Ile Ser Ile Ile Val Asn Phe Val Leu Phe Ile Ser Ile Ile Arg	
	290 295 300	
45	att ttg ctg cag aag tta aca tcc cca gat gtc ggc ggc aac gac cag	960
	Ile Leu Leu Gln Lys Leu Thr Ser Pro Asp Val Gly Gly Asn Asp Gln	
	305 310 315 320	
50	tct cag tac aag agg ctg gcc aag tcc acg ctc ctg ctt atc ccg ctg	1008
	Ser Gln Tyr Lys Arg Leu Ala Lys Ser Thr Leu Leu Leu Ile Pro Leu	
	325 330 335	
55	ttc ggc gtc cac tac atg gtg ttt gcc gtg ttt ccc atc agc atc tcc	1056
	Phe Gly Val His Tyr Met Val Phe Ala Val Phe Pro Ile Ser Ile Ser	
	340 345 350	
60	tcc aaa tac cag ata ctg ttt gag ctg tgc ctc ggc tgc ttc cag ggc	1104
	Ser Lys Tyr Gln Ile Leu Phe Glu Leu Cys Leu Gly Ser Phe Gln Gly	
	355 360 365	
65	ctg gtg gtg gcc gtc ctc tac tgt ttc ctg aac agt gag gtg cag tgc	1152
	Leu Val Val Ala Val Leu Tyr Cys Phe Leu Asn Ser Glu Val Gln Cys	
	370 375 380	

gag ctg aag cga aaa tgg cga agc cgg tgc ccg acc ccg tcc gcg agc 1200
Glu Leu Lys Arg Lys Trp Arg Ser Arg Cys Pro Thr Pro Ser Ala Ser
385 390 395 400

5 cgg gat tac agg gtc tgc ggt tcc tcc ttc tcc cac aac ggc tcg gag 1248
Arg Asp Tyr Arg Val Cys Gly Ser Ser Phe Ser His Asn Gly Ser Glu
405 410 415

10 ggc gcc ctg cag ttc cac cgc gcg tcc cga gcc cag tcc ttc ctg caa 1296
Gly Ala Leu Gln Phe His Arg Ala Ser Arg Ala Gln Ser Phe Leu Gln
420 425 430

15 acg gag acc tcg gtc atc tag 1317
Thr Glu Thr Ser Val Ile
435

20 <210> 20

25 <211> 438

30 <212> PRT

35 <213> Homo sapiens

40 <220>

45 <221> SIGNAL

50 <222> (1) .. (23)

55 <223> Potential

<220>

<221> GENE

<222> (1) .. (438)

<223> Vasoactive Intestinal Peptide Receptor 2

<220>

<221> CHAIN

<222> (24) .. (438)

<223> Vasoactive Intestinal Polypeptide Receptor 2

<220>

<221> DOMAIN

<222> (24)..(126)

<223> 1 (Potential)

<220>

<221> TRANSMEM

<222> (127)..(151)

<223> 1 (Potential)

<220>

<221> DOMAIN

<222> (152)..(158)

<223> Cytoplasmic (potential)

<220>

<221> TRANSMEM

<222> (159)..(178)

<223> 2 (Potential)

<220>

<221> DOMAIN

<222> (179)..(203)

<223> Extracellular (Potential)

<220>

<221> TRANSMEM

<222> (204)..(227)

<223> 3 (Potential)

5

<220>

<221> DOMAIN

10

<222> (228)..(240)

<223> Cytoplasmic (Potential)

15

<220>

<221> TRANSMEM

20

<222> (241)..(262)

<223> 4 (Potential)

25

<220>

<221> DOMAIN

30

<222> (263)..(279)

<223> Extracellular (Potential)

35

<220>

<221> TRANSMEM

40

<222> (280)..(303)

<223> 5 (Potential)

45

<220>

<221> DOMAIN

50

<222> (304)..(328)

<223> Cytoplasmic (Potential)

55

09/23/00 10:01
TO: JOT-00000000

5
10
15
20
25
30
35
40
45
50
55

<220>
<221> TRANSMEM
<222> (329)..(348)
<223> 6 (Potential)
<220>
<221> DOMAIN
<222> (349)..(360)
<223> Extracellular (Potential)
<220>
<221> TRANSMEM
<222> (361)..(380)
<223> 7 (Potential)
<220>
<221> DOMAIN
<222> (381)..(438)
<223> Cytoplasmic (potential)
<220>
<221> CARBOHYD
<222> (58)..(58)
<223> N-linked (GLCNAC...) (Potential)
<220>
<221> CARBOHYD
<222> (88)..(88)

<223> N-linked (GLCNAC...) (Potential)

5 <220>

<221> CARBOHYD

10 <222> (92)..(92)

<223> N-linked (GLCNAC...) (Potential)

15 <400> 20

Met Arg Thr Leu Leu Pro Pro Ala Leu Leu Thr Cys Trp Leu Leu Ala
1 5 10 15

Pro Val Asn Ser Ile His Pro Glu Cys Arg Phe His Leu Glu Ile Gln
20 25 30

Glu Glu Glu Thr Lys Cys Ala Glu Leu Leu Arg Ser Gln Thr Glu Lys
35 40 45

His Lys Ala Cys Ser Gly Val Trp Asp Asn Ile Thr Cys Trp Arg Pro
50 55 60

Ala Asn Val Gly Glu Thr Val Thr Val Pro Cys Pro Lys Val Phe Ser
65 70 75 80

Asn Phe Tyr Ser Lys Ala Gly Asn Ile Ser Lys Asn Cys Thr Ser Asp
85 90 95

Gly Trp Ser Glu Thr Phe Pro Asp Phe Val Asp Ala Cys Gly Tyr Ser
100 105 110

45 Asp Pro Glu Asp Glu Ser Lys Ile Thr Phe Tyr Ile Leu Val Lys Ala
115 120 125

50 Ile Tyr Thr Leu Gly Tyr Ser Val Ser Leu Met Ser Leu Ala Thr Gly
130 135 140

Ser Ile Ile Leu Cys Leu Phe Arg Lys Leu His Cys Thr Arg Asn Tyr
145 150 155 160

55

Ile His Leu Asn Leu Phe Leu Ser Phe Ile Leu Arg Ala Ile Ser Val
 165 170 175
 5 Leu Val Lys Asp Asp Val Leu Tyr Ser Ser Ser Gly Thr Leu His Cys
 180 185 190
 10 Pro Asp Gln Pro Ser Ser Trp Val Gly Cys Lys Leu Ser Leu Val Phe
 195 200 205
 15 Leu Gln Tyr Cys Ile Met Ala Asn Phe Phe Trp Leu Leu Val Glu Gly
 210 215 220
 Leu Tyr Leu His Thr Leu Leu Val Ala Met Leu Pro Pro Arg Arg Cys
 225 230 235 240
 20 Phe Leu Ala Tyr Leu Leu Ile Gly Trp Gly Leu Pro Thr Val Cys Ile
 245 250 255
 25 Gly Ala Trp Thr Ala Ala Arg Leu Tyr Leu Glu Asp Thr Gly Cys Trp
 260 265 270
 30 Asp Thr Asn Asp His Ser Val Pro Trp Trp Val Ile Arg Ile Pro Ile
 275 280 285
 35 Leu Ile Ser Ile Ile Val Asn Phe Val Leu Phe Ile Ser Ile Ile Arg
 290 295 300
 Ile Leu Leu Gln Lys Leu Thr Ser Pro Asp Val Gly Gly Asn Asp Gln
 305 310 315 320
 40 Ser Gln Tyr Lys Arg Leu Ala Lys Ser Thr Leu Leu Leu Ile Pro Leu
 325 330 335
 45 Phe Gly Val His Tyr Met Val Phe Ala Val Phe Pro Ile Ser Ile Ser
 340 345 350
 50 Ser Lys Tyr Gln Ile Leu Phe Glu Leu Cys Leu Gly Ser Phe Gln Gly
 355 360 365
 55 Leu Val Val Ala Val Leu Tyr Cys Phe Leu Asn Ser Glu Val Gln Cys
 370 375 380

Glu Leu Lys Arg Lys Trp Arg Ser Arg Cys Pro Thr Pro Ser Ala Ser
385 390 395 400

5 Arg Asp Tyr Arg Val Cys Gly Ser Ser Phe Ser Arg Asn Gly Ser Glu
405 410 415

10 Gly Ala Leu Gln Phe His Arg Gly Ser Arg Ala Gln Ser Phe Leu Gln
420 425 430

Thr Glu Thr Ser Val Ile
435

15

<210> 21

<211> 889

<212> DNA

<213> Homo sapiens

<220>

<221> REFERENCE

<222> (1) .. (889)

<223> Li, Y.S., Milner, P.G., Chauhan, A.K., Watson, M.A., Hoffman, R.M
., Kodner, C.M., Milbrandt, J. and Deuel, T.F. TITLE Cloning and
expression of a developmentally regulated protein that induces mi
togenic and neurite outgrowth activity JOURNAL Science 250 (4988)
, 1690-1694 (1990)

35

<220>

<221> Gene

<222> (1) .. (889)

<223> Pleiotrophin

50

<220>

<221> CDS

55

<222> (252) .. (758)

<223>

5 <400> 21
gtcaaaggca ggatcaggtt cccgccttc cagtccaaaa atcccgccaa gagagcccca 60
gagcagagga aaatccaaag tggagagagg ggaagaaaga gaccagttag tcatccgtcc 120
10 agaaggcggg gagagcagca gcggcccaag caggagctgc agcgagccgg gtacctggac 180
tcagcggttag caacctcgcc ccttgcaaca aaggcagact gagcgccaga gaggacgttt 240
ccaactcaaa a atg cag gct caa cag tac cag cag cag cgt cga aaa ttt 290
15 Met Gln Ala Gln Gln Tyr Gln Gln Gln Arg Lys Phe
1 5 10
gca gct gcc ttc ttg gca ttc att ttc ata ctg gca gct gtg gat act 338
Ala Ala Ala Phe Leu Ala Phe Ile Phe Ile Leu Ala Ala Val Asp Thr
15 20 25
gct gaa gca ggg aag aaa gag aaa cca gaa aaa aaa gtg aag aag tct 386
Ala Glu Ala Gly Lys Lys Glu Lys Pro Glu Lys Lys Val Lys Lys Ser
30 35 40 45
gac tgt gga gaa tgg cag tgg agt gtg tgt gtg ccc acc agt gga gac 434
Asp Cys Gly Glu Trp Gln Trp Ser Val Cys Val Pro Thr Ser Gly Asp
50 55 60
tgt ggg ctg ggc aca cgg gag ggc act cgg act gga gct gag tgc aag 482
Cys Gly Leu Gly Thr Arg Glu Gly Thr Arg Thr Gly Ala Glu Cys Lys
65 70 75
caa acc atg aag acc cag aga tgt aag atc ccc tgc aac tgg aag aag 530
Gln Thr Met Lys Thr Gln Arg Cys Lys Ile Pro Cys Asn Trp Lys Lys
80 85 90
caa ttt ggc gcg gag tgc aaa tac cag ttc cag gcc tgg gga gaa tgt 578
Gln Phe Gly Ala Glu Cys Lys Tyr Gln Phe Gln Ala Trp Gly Glu Cys
95 100 105
gac ctg aac aca gcc ctg aag acc aga act gga agt ctg aag cga gcc 626
Asp Leu Asn Thr Ala Leu Lys Thr Arg Thr Gly Ser Leu Lys Arg Ala
110 115 120 125
45 ctg cac aat gcc gaa tgc cag aag act gtc acc atc tcc aag ccc tgt 674
Leu His Asn Ala Glu Cys Gln Lys Thr Val Thr Ile Ser Lys Pro Cys
130 135 140
50 ggc aaa ctg acc aag ccc aaa cct caa gca gaa tct aag aag aag aaa 722
Gly Lys Leu Thr Lys Pro Lys Pro Gln Ala Glu Ser Lys Lys Lys Lys
145 150 155
aag gaa ggc aag aaa cag gag aag atg ctg gat taa aagatgtcac 768
55 Lys Glu Gly Lys Lys Gln Glu Lys Met Leu Asp
160 165

ctgtggaaca taaaaaggac atcagcaaac aggatcagtt aactattgca tttatatgta 828

ccgtaggctt tgtattcaaa aattatctat agctaagtac acaataagca aaaacaaaaa 888

g 889

<210> 22

<211> 168

<212> PRT

<213> Homo sapiens

<220>

<221> SIGNAL

<222> (1) .. (32)

<223>

<220>

<221> GENE

<222> (1) .. (168)

<223> Pleiotrophin

<220>

<221> CHAIN

<222> (33) .. (168)

<223> Pleiotrophin

<220>

<221> DISULFID

<222> (47) .. (76)

<223> By similarity.

5 <220>
<221> DISULFID
<222> (55)..(85)
10 <223> By similarity.

15 <220>
<221> DISULFID
<222> (62)..(89)
20 <223> By similarity.

25 <220>
<221> DISULFID
<222> (99)..(131)
30 <223> By similarity.

35 <220>
<221> DISULFID
<222> (109)..(141)
40 <223> By similarity.

45 <400> 22
Met Gln Ala Gln Gln Tyr Gln Gln Gln Arg Arg Lys Phe Ala Ala Ala
1 5 10 15
50 Phe Leu Ala Phe Ile Phe Ile Leu Ala Ala Val Asp Thr Ala Glu Ala
20 25 30
55 Gly Lys Lys Glu Lys Pro Glu Lys Lys Val Lys Lys Ser Asp Cys Gly
35 40 45

Glu Trp Gln Trp Ser Val Cys Val Pro Thr Ser Gly Asp Cys Gly Leu
50 55 60

5
Gly Thr Arg Glu Gly Thr Arg Thr Gly Ala Glu Cys Lys Gln Thr Met
65 70 75 80

10
Lys Thr Gln Arg Cys Lys Ile Pro Cys Asn Trp Lys Lys Gln Phe Gly
85 90 95

15
Ala Glu Cys Lys Tyr Gln Phe Gln Ala Trp Gly Glu Cys Asp Leu Asn
100 105 110

20
Thr Ala Leu Lys Thr Arg Thr Gly Ser Leu Lys Arg Ala Leu His Asn
115 120 125

25
Ala Glu Cys Gln Lys Thr Val Thr Ile Ser Lys Pro Cys Gly Lys Leu
130 135 140

30
Thr Lys Pro Lys Pro Gln Ala Glu Ser Lys Lys Lys Lys Lys Glu Gly
145 150 155 160

35
Lys Lys Gln Glu Lys Met Leu Asp
165

<210> 23

<211> 3143

<212> DNA

<213> Homo sapiens

<220>

<221> REFERENCE

<222> (1)..(3143)

50 <223> Crosby, A.H., Edwards, S.J., Murray, J.C. and Dixon, M.J. TITLE G
enomic organization of the human osteopontin gene: exclusion of t
he locus from a causative role in the pathogenesis of dentinogene
sis imperfecta type II JOURNAL Genomics 27 (1), 155-160 (1995)

$\langle 220 \rangle$

<221> Gene

5 <222> (1) .. (3143)

<223> Osteopontin

10

 $\langle 220 \rangle$

<221> CDS

15 $\langle 222 \rangle$ (374) . . (427)

<223>

20

<220>

<221> CDS

25 <222> (537) .. (575)

<223>

30

 $\langle 220 \rangle$

<221> CDS

35 <222> (884) . . (964)

<223>

40

 $\langle 220 \rangle$

<221> CDS

45 <222> (1232) .. (1273)

<223>

50

<220>

<221> CDS

55 <222> (1597) .. (1920)

<223>

5 <220>

<221> CDS

<222> (2305) .. (2709)

10 <223>

15 <400> 33

ggggaagtgt gggagcaggt gggctgggca gtggcagaaa cctgatgaca caatctcgcc 60

gcctccctgt gttggtggag gatgtctgca gcagcattta aattctggga gggcttggtt 120

gtcagcagca gcaggaggag gcagagacag catcgtcggg accagactcg tctcaggcca 180

gttgcagcct tctcagccaa acgccgacca aggtacagct tcagtttgct actgggttgt 240

gcattcagct gaatttcatg gggaagtcca aattctaagg aaaaaaatgt ggtagtataa 300

aaaggtatca ctgttgtaac ctatgaagat gtcagctatt cctttgaaat attttgcagg 360

aaaactcact acc atg aga att gca gtg att tgc ttt tgc ctc cta ggc 409
Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly
1 5 10

atc acc tgt gcc ata cca gtgagtacag ttgcatctta aagaaaattc 457
Ile Thr Cys Ala Ile Pro
15

ctgaaaataa ctgaattgtg tgcttccatg tgctaggagg acattcttgt aatctttctt 517

catcttttct gtttctaag gtt aaa cag gct gat tct gga agt tct gag gaa 569
Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu
20 25

aag cag gtaagcatct tttatgtttt tatatagtta aatcatttac tcaattatgg 625
Lys Gln
30

cgagaggtgc aagaaacgta tttgctgcga tcaaagagt tcatatttgt aaagcaattt 685

gaaagagtgc ctagcccaaca gtaagtgcata cataagagtt tgtaaataa atctgcaaaa 745

aaaaaaaaaa ttacaaaaag gtacctaagg gtccgggtga ctatatgctt ccatcaagac 805

tagtgaagaa tgggtgtttt ttccattcat ccctacattt ctttttttaa taatgataaa 865

catgcaactt tttttag ctt tac aac aaa tac cca gat gct gtg gcc aca 916
Leu Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr
35 40

5
10
15
20
25
30
35
40
45
50
55

tgg cta aac cct gac cca tct cag aag cag aat ctc cta gcc cca cag 964
Trp Leu Asn Pro Asp Pro Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln
45 50 55

gtattttttaa acttttcata attaaactac agtgatgaaa gatagccaca ctcaggccat 1024
ttgggctgct cagatgaatc ctgccctgcc tgctggcaaa catgtgctta ggacattgac 1084
tgatctgccca tgttggcttc tctctgtgtt aagccatcca cagatgaggc tgaaaaataa 1144
aaactgcttt ggattaaaaa ggtaacttt tgaataaaaa agctaggcat gtgtgatgag 1204
cactaacacg tgccattcct tcttcag aat gct gtg tcc tct gaa gaa acc aat 1258
Asn Ala Val Ser Ser Glu Glu Thr Asn
60 65

gac ttt aaa caa gag gtaagttctc attttcaatc agaggcccat catgccttga 1313
Asp Phe Lys Gln Glu
70

agagatgaaa gaaggcattg cctggattct cttctgatga aatttcatta gcaagttttc 1373
cagctaattg gcagtctaaa acttgctcat aaataaaaaca tgtattttact aaatatcaga 1433
aataactagg ttcctcgat aacctaaga ccatgggatg tactgtgaat gcaaagattc 1493
tgaaactaaa taaaaagaaa gatagtaaaa gactaatgtg ctataaaggc taagggaataa 1553
taaaaaccca tatattaatt ttcccggcca tcttaatttt cag acc ctt cca agt 1608
Thr Leu Pro Ser
75

aag tcc aac gaa agc cat gac cac atg gat gat atg gat gat gaa gat 1656
Lys Ser Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp
80 85 90

gat gat gac cat gtg gac agc cag gac tcc att gac tcg aac gac tct 1704
Asp Asp Asp His Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser
95 100 105

gat gat gta gat gac act gat gat tct cac cag tct gat gag tct cac 1752
Asp Asp Val Asp Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His
110 115 120

cat tct gat gaa tct gat gaa ctg gtc act gat ttt ccc acg gac ctg 1800
His Ser Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu
125 130 135 140

cca gca acc gaa gtt ttc act cca gtt gtc ccc aca gta gac aca tat 1848
Pro Ala Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr
145 150 155

gat ggc cga ggt gat agt gtg gtt tat gga ctg agg tca aaa tct aag 1896
Asp Gly Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys
160 165 170

5 aag ttt cgc aga cct gac atc cag gtaaattcctt taacagacac acctgatggt 1950
Lys Phe Arg Arg Pro Asp Ile Gln
175 180

10 tctgactagc gctcaagtct aggaaccac agtttgcata ttcattcatt cattcatcca 2010
ttcattcatc cattcagcaa gaattcattc atattctact ttatgaccat tgaatacaaaa 2070

15 tctttttctg cttggcgggtt tttgtaagtc tacataattt ctctctagat ttgattctca 2130
aacacaattc tacttttttga aatcctggat caaagtaaca tgctagtatt atttcagcca 2190
gatttagaca atttttagta taagatgacc taaaagctag agagtggaaa aggattacca 2250
tattcccatc cctagccgtt catataatta ttcttcattt gtgccgtgat tcag tac 2307
Tyr

20 cct gat gct aca gac gag gac atc acc tca cac atg gaa agc gag gag 2355
Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His Met Glu Ser Glu Glu
185 190 195

25 ttg aat ggt gca tac aag gcc atc ccc gtt gcc cag gac ctg aac gcg 2403
Leu Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala Gln Asp Leu Asn Ala
200 205 210

30 cct tct gat tgg gac agc cgt ggg aag gac agt tat gaa acg agt cag 2451
Pro Ser Asp Trp Asp Ser Arg Gly Lys Asp Ser Tyr Glu Thr Ser Gln
215 220 225

35 ctg gat gac cag agt gct gaa acc cac agc cac aag cag tcc aga tta 2499
Leu Asp Asp Gln Ser Ala Glu Thr His Ser His Lys Gln Ser Arg Leu
230 235 240 245

40 tat aag cgg aaa gcc aat gat gag agc aat gag cat tcc gat gtg att 2547
Tyr Lys Arg Lys Ala Asn Asp Glu Ser Asn Glu His Ser Asp Val Ile
250 255 260

45 gat agt cag gaa ctt tcc aaa gtc agc cgt gaa ttc cac agc cat gaa 2595
Asp Ser Gln Glu Leu Ser Lys Val Ser Arg Glu Phe His Ser His Glu
265 270 275

50 ttt cac agc cat gaa gat atg ctg gtt gta gac ccc aaa agt aag gaa 2643
Phe His Ser His Glu Asp Met Leu Val Val Asp Pro Lys Ser Lys Glu
280 285 290

55 gaa gat aaa cac ctg aaa ttt cgt att tct cat gaa tta gat agt gca 2691
Glu Asp Lys His Leu Lys Phe Arg Ile Ser His Glu Leu Asp Ser Ala
295 300 305

tct tct gag gtc aat taa aaggagaaaa aatacaattt ctcactttgc 2739
Ser Ser Glu Val Asn
310

atttagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgcttcttt 2799

ctcagtttat tggttgaatg tgtatctatt tgagtctgga aataactaat gtgtttgata 2859
attagtttag tttgtggctt catggaaact ccctgtaaac aaaagcttca gggttatgtc 2919
5 tatgttcatt ctatagaaga aatgcaaact atcactgtat tttaatatth gttattctct 2979
catgaataga aatttatgta gaagcaaaca aaatactttt acccacttaa aaagagaata 3039
10 taacatttta tgtcactata atcttttgth ttttaagtha gtgtatattt tgttgtgatt 3099
atcttttgth gtgtgaataa atcttttatc ttgaatgtaa taag 3143

15 <210> 24
<211> 314
<212> PRT
20 <213> Homo sapiens

25 <220>
<221> SIGNAL
<222> (1)..(16)
30 <223> Potential

35 <220>
<221> Gene
<222> (1)..(314)
40 <223> Osteopontin

45 <220>
<221> CHAIN
<222> (17)..(314)
50 <223> Osteopontin

55 <220>

5 <221> SITE
<222> (159)..(161)
5 <223> Cell attachment site

10 <220>
<221> CARBOHYD
<222> (79)..(79)
15 <223> N-linked (GLCNAC...) (Potential)

20 <220>
<221> CARBOHYD
<222> (106)..(106)
25 <223> N-linked (GLCNAC...) (Potential)

30 <220>
<221> VARSPLIC
<222> (31)..(57)
35 <223> Missing (In Isoform C)

40 <220>
<221> VARSPLIC
<222> (58)..(71)
45 <223> Missing (In Isoform B)

50 <400> 24
Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala
1 5 10 15
55 Ile Pro Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu Lys Gln Leu
20 25 30

5 Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu Asn Pro Asp Pro
35 40 45

10 Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Asn Ala Val Ser Ser Glu
50 55 60

15 Glu Thr Asn Asp Phe Lys Gln Glu Thr Leu Pro Ser Lys Ser Asn Glu
65 70 75 80

20 Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp Asp His
85 90 95

25 Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp Val Asp
100 105 110

30 Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His His Ser Asp Glu
115 120 125

35 Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu Pro Ala Thr Glu
130 135 140

40 Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr Asp Gly Arg Gly
145 150 155 160

45 Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe Arg Arg
165 170 175

50 Pro Asp Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His
180 185 190

55 Met Glu Ser Glu Glu Leu Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala
195 200 205

60 Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp Ser Arg Gly Lys Asp Ser
210 215 220

65 Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His Ser His
225 230 235 240

70 Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser Asn Glu
245 250 255

His Ser Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser Arg Glu
260 265 270

Phe His Ser His Glu Phe His Ser His Glu Asp Met Leu Val Val Asp
275 280 285

Pro Lys Ser Lys Glu Glu Asp Lys His Leu Lys Phe Arg Ile Ser His
290 295 300

Glu Leu Asp Ser Ala Ser Ser Glu Val Asn
305 310

<210> 25

<211> 259

<212> PRT

<213> Homo sapiens

<220>

<221> Gene

<222> (1)..(259)

<223> Carbonic Anhydrase domain of human carbonic anhydrase III

<400> 41

Ala Lys Glu Trp Gly Tyr Ala Ser His Asn Gly Pro Asp His Trp His
1 5 10 15

Glu Leu Phe Pro Asn Ala Lys Gly Glu Asn Gln Ser Pro Ile Glu Leu
20 25 30

His Thr Lys Asp Ile Arg His Asp Pro Ser Leu Gln Pro Trp Ser Val
35 40 45

Ser Tyr Asp Gly Gly Ser Ala Lys Thr Ile Leu Asn Asn Gly Lys Thr
50 55 60

5 Cys Arg Val Val Phe Asp Asp Thr Tyr Asp Arg Ser Met Leu Arg Gly
 65 70 75 80
 Gly Pro Leu Pro Gly Pro Tyr Arg Leu Arg Gln Phe His Leu His Trp
 85 90 95
 10 Gly Ser Ser Asp Asp His Gly Ser Glu His Thr Val Asp Gly Val Lys
 100 105 110
 Tyr Ala Ala Glu Leu His Leu Val His Trp Asn Pro Lys Tyr Asn Thr
 115 120 125
 15 Phe Lys Glu Ala Leu Lys Gln Arg Asp Gly Ile Ala Val Ile Gly Ile
 130 135 140
 Phe Leu Lys Ile Gly His Glu Asn Gly Glu Phe Gln Ile Phe Leu Asp
 145 150 155 160
 20 Ala Leu Asp Lys Ile Lys Thr Lys Gly Lys Glu Ala Pro Phe Thr Lys
 165 170 175
 Phe Asp Pro Ser Cys Leu Phe Pro Ala Cys Arg Asp Tyr Trp Thr Tyr
 180 185 190
 25 Gln Gly Ser Phe Thr Thr Pro Pro Cys Glu Glu Cys Ile Val Trp Leu
 195 200 205
 30 Leu Leu Lys Glu Pro Met Thr Val Ser Ser Asp Gln Met Ala Lys Leu
 210 215 220
 35 Arg Ser Leu Leu Ser Ser Ala Glu Asn Glu Pro Pro Val Pro Leu Val
 225 230 235 240
 40 Ser Asn Trp Arg Pro Pro Gln Pro Ile Asn Asn Arg Val Val Arg Ala
 245 250 255
 45 Ser Phe Lys
 50
 <210> 26
 55 <211> 260

<212> PRT

<213> Homo sapiens

<220>

<221> Gene

<222> (1) .. (260)

<223> Carbonic anhydrase domain of human carbonic anhydrase I

<400> 26

Ala Ser Pro Asp Trp Gly Tyr Asp Asp Lys Asn Gly Pro Glu Gln Trp
1 5 10 15

Ser Lys Leu Tyr Pro Ile Ala Asn Gly Asn Asn Gln Ser Pro Val Asp
20 25 30

Ile Lys Thr Ser Glu Thr Lys His Asp Thr Ser Leu Lys Pro Ile Ser
35 40 45

Val Ser Tyr Asn Pro Ala Thr Ala Lys Glu Ile Ile Asn Val Gly His
50 55 60

Ser Phe His Val Asn Phe Glu Asp Asn Asp Asn Arg Ser Val Leu Lys
65 70 75 80

Gly Gly Pro Phe Ser Asp Ser Tyr Arg Leu Phe Gln Phe His Phe His
85 90 95

Trp Gly Ser Thr Asn Glu His Gly Ser Glu His Thr Val Asp Gly Val
100 105 110

Lys Tyr Ser Ala Glu Leu His Val Ala His Trp Asn Ser Ala Lys Tyr
115 120 125

Ser Ser Leu Ala Glu Ala Ala Ser Lys Ala Asp Gly Leu Ala Val Ile
130 135 140

Gly Val Leu Met Lys Val Gly Glu Ala Asn Pro Lys Leu Gln Lys Val
145 150 155 160

5

Leu Asp Ala Leu Gln Ala Ile Lys Thr Lys Gly Lys Arg Ala Pro Phe
165 170 175

10

Thr Asn Phe Asp Pro Ser Thr Leu Leu Pro Ser Ser Leu Asp Phe Trp
180 185 190

15

Thr Tyr Pro Gly Ser Leu Thr His Pro Pro Leu Tyr Glu Ser Val Thr
195 200 205

20

Trp Ile Ile Cys Lys Glu Ser Ile Ser Val Ser Ser Glu Gln Leu Ala
210 215 220

25

Gln Phe Arg Ser Leu Leu Ser Asn Val Glu Gly Asp Asn Ala Val Pro
225 230 235 240

30

Met Gln His Asn Asn Arg Pro Thr Gln Pro Leu Lys Gly Arg Thr Val
245 250 255

35

Arg Ala Ser Phe
260

40

<210> 27

<211> 337

<212> PRT

<213> Homo sapiens

45

<220>

<221> Gene

<222> (1)..(337)

<223> Carbonic anhydrase domain of human carbonic anhydrase VIX

50

<400> 27

55

Met Leu Phe Ser Ala Leu Leu Leu Glu Val Ile Trp Ile Leu Ala Ala
1 5 10 15

Asp Gly Gly Gln His Trp Thr Tyr Glu Gly Pro His Gly Gln Asp His
20 25 30

5 Trp Pro Ala Ser Tyr Pro Glu Cys Gly Asn Asn Ala Gln Ser Pro Ile
35 40 45

10 Asp Ile Gln Thr Asp Ser Val Thr Phe Asp Pro Asp Leu Pro Ala Leu
50 55 60

Gln Pro His Gly Tyr Asp Gln Pro Gly Thr Glu Pro Leu Asp Leu His
65 70 75 80

15 Asn Asn Gly His Thr Val Gln Leu Ser Leu Pro Ser Thr Leu Tyr Leu
85 90 95

Gly Gly Leu Pro Arg Lys Tyr Val Ala Ala Gln Leu His Leu His Trp
100 105 110

25 Gly Gln Lys Gly Ser Pro Gly Gly Ser Glu His Gln Ile Asn Ser Glu
115 120 125

Ala Thr Phe Ala Glu Leu His Ile Val His Tyr Asp Ser Asp Ser Tyr
130 135 140

Asp Ser Leu Ser Glu Ala Ala Glu Arg Pro Gln Gly Leu Ala Val Leu
145 150 155 160

Gly Ile Leu Ile Glu Val Gly Glu Thr Lys Asn Ile Ala Tyr Glu His
165 170 175

40 Ile Leu Ser His Leu His Glu Val Arg His Lys Asp Gln Lys Thr Ser
180 185 190

45 Val Pro Pro Phe Asn Leu Arg Glu Leu Leu Pro Lys Gln Leu Gly Gln
195 200 205

50 Tyr Phe Arg Tyr Asn Gly Ser Leu Thr Thr Pro Pro Cys Tyr Gln Ser
210 215 220

Val Leu Trp Thr Val Phe Tyr Arg Arg Ser Gln Ile Ser Met Glu Gln
225 230 235 240

Leu Glu Lys Leu Gln Gly Thr Leu Phe Ser Thr Glu Glu Glu Pro Ser
245 250 255

5 Lys Leu Leu Val Gln Asn Tyr Arg Ala Leu Gln Pro Leu Asn Gln Arg
260 265 270

10 Met Val Phe Ala Ser Phe Ile Gln Ala Gly Ser Ser Tyr Thr Thr Gly
275 280 285

15 Glu Met Leu Ser Leu Gly Val Gly Ile Leu Val Gly Cys Leu Cys Leu
290 295 300

Leu Leu Ala Val Tyr Phe Ile Ala Arg Lys Ile Arg Lys Lys Arg Leu
305 310 315 320

20 Glu Asn Arg Lys Ser Val Val Phe Thr Ser Ala Gln Ala Thr Thr Glu
325 330 335

25 Ala

<210> 28

<211> 22

<212> DNA

35 <213> Artificial sequence

<220>

40 <223> Primer

<400> 28

cagcagttgg atggaagagg ac

22

<210> 29

<211> 22

50 <212> DNA

<213> Artificial sequence

55

5 <220>
<223> Primer
<400> 29
cactgagatt ctggcactat tc
10 <210> 30
<211> 21
<212> DNA
15 <213> Artificial sequence
20 <220>
<223> Primer
<400> 30
aacaattcca gggctctcact c
25 <210> 31
<211> 21
<212> DNA
30 <213> Artificial sequence
35 <220>
<223> Primer
40 <400> 31
ttgactggct caggagtata g
45 <210> 32
<211> 21
<212> DNA
50 <213> Artificial sequence
55 <220>

22

21

21

<223> Primer

<400> 32

ctgataatga gggctcccaa c

21

<210> 33

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 33

ctctgcactt cctggtaaaa ctct

24

<210> 34

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 34

cagcagttgg atggaagagg ac

22

<210> 35

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 35
ctctgcactt cctggtaaaa ctct

24

5

TTTTT-00000000